Universidad Autónoma de Madrid

Facultad de Ciencias

Departamento de Biología Molecular

Functional analysis of two Arabidopsis SET domain-containing proteins (SDG7 and SDG24) during development

Carla Dianela Méndez

Madrid, 2018

Departamento de Biología Molecular

Facultad de Ciencias

Universidad Autónoma de Madrid

Programa de Doctorado de Biociencias Moleculares

Functional analysis of two Arabidopsis SET domain-containing proteins (SDG7 and SDG24) during development

Carla Dianela Méndez
Licenciada en Biología

Dirigida por:
Dr. Crisanto Gutiérrez

Lugar de Realización:
Centro de Biología Molecular Severo Ochoa
Consejo Superior de investigaciones Científicas
(CBMSO-CSIC)

Madrid, 2018

El Dr. Crisanto Gutiérrez, Profesor de Investigación CSIC en el Centro de Biología Molecular Severo Ochoa

CERTIFICA

Que la Tesis Doctoral titulada "Functional analysis of two Arabidopsis SET domain-containing proteins (SDG7 and SDG24) during development" ha sido realizada en el Centro de Biología Molecular Severo Ochoa y tutelada en el Departamento de Biología Molecular de la Universidad Autónoma de Madrid.

El trabajo realizado por Doña Carla Dianela Méndez reúne todas las condiciones requeridas por la legislación vigente, así como la originalidad y calidad científica para poder ser presentada y defendida con el fin de optar al grado de Doctor.

Este trabajo cuenta con el visto bueno del Dr. Luis Blanco Dávila, tutor académico en el Departamento de Biología Molecular de la Universidad Autónoma de Madrid.

Y para que conste donde proceda, firmo el presente certificado
Madrid a 19 de Octubre de 2017

Dr. Crisanto Gutiérrez
Director de la Tesis

El Dr. Luis Blanco, Profesor de Investigación CSIC en el Centro de Biología Molecular Severo Ochoa

CERTIFICA

Que la Tesis Doctoral titulada "Functional analysis of two Arabidopsis SET domain-containing proteins (SDG7 and SDG24) during development" ha sido realizada en el Centro de Biología Molecular Severo Ochoa y tutelada en el Departamento de Biología Molecular de la Universidad Autónoma de Madrid.

El trabajo realizado por Doña Carla Dianela Méndez reúne todas las condiciones requeridas por la legislación vigente, así como la originalidad y calidad científica para poder ser presentada y defendida con el fin de optar al grado de Doctor.

Este trabajo cuenta con mi visto bueno como tutor académico en el Departamento de Biología Molecular de la Universidad Autónoma de Madrid.

Y para que conste donde proceda, firmo el presente certificado
Madrid a 19 de Octubre de 2017

Dr. Luis Blanco
Tutor de la Tesis

A Venezuela y España

A mis padres

A Daniel

“Post nubila, Phoebus”

Proverbio latino

“Change will not come if we wait for some other person or some other time. We are the ones we've been waiting for. We are the change that we seek”

Barack Obama

Table of contents

Acknowledgments/Agradecimientos

Había una vez, una chica de Venezuela, que es italiana, y que se fue desde Valencia a Madrid a hacer el doctorado. Después de varios años de experimentos, desea hacer unos agradecimientos muy convenientes. Primeramente agradezco al reino de **España** y al **MECD** por otorgarme la FPU, la cual me dio la oportunidad de cursar los estudios de doctorado. Agradezco también al honorable Crisanto, **CG_ath**, el director, quien con sus buenas ideas, y su amplia sabiduría condujo este proyecto, proporcionó una perspectiva de que allá del núcleo esta la cromatina, abrió un horizonte al avistamiento de nucléolos, y el cual ha sido un ejemplo de trabajo constante. A **Béné** (ring.. ring: …esta Bénédicte?), a esa gran mujer investigadora y luchadora, la mejor enciclopedia (mejor que la wiki y muchos papers), que sirvió de apoyo para muchas discusiones (científicas y políticas), por su disponibilidad, sus enseñanzas, y por sus trucos experimentales totalmente fiables (muy valorados y apreciados), gracias. A **Joana** por sus invaluables lecciones científicas, los ORIS y sus consejos, por su ayuda durante gran parte de este trabajo. A **Zaida**, por su tiempo, sus protocolos, por haber visto al nucléolo en el microscopio, por brindar con Don J. Cuervo en las buenas y en las malas, por recordarnos el verano todo el año. A **María FM**, una gran compañera, amiga e investigadora, que me enseñó todo lo que sabia de raíces, sobre microscopía, sobre Arabidopsis, sobre la vida. A **Sofía** Otero, por tener esa chispa científica motivadora, las ganas de saber, de enseñar, de estar siempre alegre… chachis!. A **Sofia** Madeira, por su generosidad, su compromiso y su dedicación al trabajo, y sobre todo por su sencillez y calidez, y también a **Ricardo**, *muito obrigada*! A **Victoria**, nuestra Viki, por su gran trabajo de apoyo con las plantas y sobre todo por sus grandes historias de viajes por el mundo, que nos daban para soñar con el conocer otros lares. A **Iluminada**, la luz del sol en las tardes, en las noches, la levanta almas, la arranca sonrisas, nuestra madre a medio tiempo, te debemos las ganas de terminar el día (o la noche) con una sonrisa. A **Adriana**, por enseñarme a cruzar flores y a **Irene** por ser una gran maestra del machacado con nitrógeno. A las chicas del té: **Gloria, Rosa** y **Charo**, que entre infusiones aromáticas discutimos curiosidades, cosas científicas y protocolos que funcionasen. A **María Gómez** y su laboratorio, por sus apreciados comentarios e ideas en los *meetings* conjuntos. Al todo el servicio de **Microscopia** del CBMSO, en especial a **Carmen** por enseñarme el Nikon y alguno de sus trucos. A las chicas del servicio de citometría, **Berta** y **Silvia**, por su gran dedicación, paciencia y compromiso de calidad hacia el trabajo que hacen. Al servicio de **Instrumentación** del CBMSO por prestarme tantas cosas y hacer posible varios experimentos caseros. Al servicio de Fermentación, en especial a **Dionisio**, que me aconsejó en todo lo que pudo con los temas de proteínas y bacterias. A **Juan,** por

ayudarnos con tantos medios. To **Martina** Dvořáčková and **Jiří** Fajkus**,** for their knowledge in nucleoli, their partnership to the ASHHs project, and for unravel new excitement things. A **Alfredo**, por enseñarme tantas cosas de laboratorio y apoyarme en la búsqueda del doctorado, por ser un ejemplo a seguir, aun desde Umea. Al profe **Carlos**, por apuntarme el camino a España cuando terminé mi licenciatura. A **Paco Madueño**, por darme la oportunidad de iniciarme en lo molecular de las plantas. A Daniel Arismendi (**Danielito**) por ser nuestra familia en el exilio y por compartir tantas noches de doctorado. A **Claudio**, **Carla** y **Alonso**, esa otra familia que a través de mis años de estudios nos hemos apoyado en seguir nuestros sueños. A **Daniela** y **Raquel**, mis compañeras y amigas, por darme aliento y apoyo cuando más lo necesitaba. A **Evaristo** e **Isabel**, mis padres, por darlo todo para que yo estuviera aquí, por enseñarme el método científico, por sus sacrificios, y por su amor. A **Daniel M. Tong**, por su ejemplo como profesional, por su entusiasmo, su apoyo emocional y sus formas de resolver las crisis científicas existenciales, por todo lo que ha dado sin pedir nada a cambio, y por mucho más. Y a las personas y o entidades que no he nombrado, mis disculpas, han sido todas y cada una excepcionales para esta etapa de formación. Y finalmente doy gracias a **Madrid**, esta gran ciudad que ha sido mi hogar.

ABBREVIATIONS

A
ACT2: ACTIN-2
AG: AGAMOUS
AGL: AGAMOUS LIKE
AGO: ARGONAUTE
AMS: ABORTED MICROSPORES
ASH1: ABSENT, SMALL, OR HOMEOTIC DISCS 1
ASHH3: ARABIDOPSIS ASH1 HOMOLOG 3
ASHH4: ARABIDOPSIS ASH1 HOMOLOG 4
ATXR: ARABIDOPSIS TXR RELATED
AWS: associated with SET

B
BCA: Pierce's bicinchoninic acid assay
BRCA1: BREAST CANCER SUSCEPTIBILITY-1
BSA: bovine serum albumin
Bur-0: Burren-0 ecotype

C
CAF-1: Chromatin assembling factor 1
CaMV: Cauliflower mosaic virus
Can-0: Canary Islands -0 ecotype
CCS52: CELL CYCLE SWITCH 52
CCH: nucleotides where H= A, C, OR T
CDC6: CELL DIVISION CYCLE 6
CDK: CYCLIN-DEPENDENT KINASES
cDNA: complementary DNA
CDS: coding sequence
CDT1: CHROMATIN LICENSING AND DNA REPLICATION FACTOR 1
CFP: cyan fluorescent protein
CHH: where H=A, T or C
ChiP: Chromatin immunoprecipitation
CHR19: CHROMATIN REMODELER-19
CLF: CURLYLEAF
Col-0: Arabidopsis Columbia-0 ecotype
Ct-1: Catania-1 ecotype
CTAB: Cetyl trimethylammonium bromide
CYC: CYCLIN
CYCB1: CYCLIN-DEPENDENT PROTEIN KINASE 1

D
DAPI: 4′, 6-diamidino-2-phenylindole
DDM1: DECREASED DNA METHYLATION-1
DFC: nucleolar dense fibrillar component
DME: DEMETER
DNA: Deoxyribonucleic acid
dps: days post sowing
DRM2: DOMAINS REARRANGED METHYLASE-2
DSB: double strand breaks
DTT: Dithiothreitol
DF: differentiation zone

E
E2F: E2 FACTOR
EDA3: EMBRYO SAC DEVELOPMENT ARREST-3
Edi-0: Edinburgh-0 ecotype
EdU: 5-Ethynyl-2′-deoxyuridine
EMF2: EMBRYONIC FLOWER-2
ERC: extrachromosomal rDNA circle
ERF115: ETHYLENE RESPONSE FACTOR-115
ESCR: equal sister chromatid recombination
ETS: external transcribed spacer
E(z): ENHANCER OF ZESTE
EZ: elongation zone

F
F1: filial generation 1
FAS1: FASCIATA-1
FCs: nucleolar fibrillar centers
PHE1: PHERES-1
FIE: FERTILIZATION-INDEPENDENT ENDOSPERM
FISH: fluorescent *in situ* hybridization
FIS2: FERTILIZATION INDEPENDENT SEED-2
FLC: FLOWERING LOCUS C
FM: floral meristem
FOB1: REPLICATION FORK BLOCKING-1

G
G1: post-mitotic gap phase
G2: post-synthetic gap phase
GAPDH: GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE
GC: nucleolar granular component
GUS: β- GLUCURONIDASE
GFP: green fluorescent protein
GST: GLUTATHIONE S-TRANSFERASE

H
H1: histone H1
H2A: histone H2A
H2B: histone H2B
H3: histone H3
H4: histone H4
HAT: histone acetyltransferases
HDA6: HISTONE DEACETYLACE 6
HDAC: histone deacetylases
HDM: histone demethylases
Hi-0: Hilversum-0 ecotype
HIS6X: 6 residues of histidine
HMT: histone methyltransferases
HSP40: HEAT SHOCK PROTEIN-40
HTR5: HISTONE THREE RELATED 5

I
IM: inflorescent meristem

K
K: lysine residue
KYP: KRYPTONITE
KD: knock-out mutant
KO: knock-down mutant

L
LD: long day
LDL1: LYSINE SPECIFIC DEMETHYLASE LIKE-1
Ler-0: Landsberg erecta-0 ecotype
LRs: lateral roots
LRP: lateral root primordia

M
M: mitosis
MBD6: METHYLCYTOSINE-BINDING DOMAIN -6
MCM: MINICHROMOSOME MAINTENANCE COMPLEX COMPONENT
MEA: MEDEA
MET1: DNA METHYLTRANSFERASE-1
me1: 1 methyl group
miRNA: microRNA
mir171a: passenger strand microrna MIR171A
mRFP: monomeric red fluorescent protein
MSA: mitosis-specific activator element
MYB: MYB PROTO-ONCOGENE
MYB3R1: MYB3 PROTEIN CONTAINING THE REPEAT R1

N
NCED3: NINE CIS EPOXYCAROTENOID DIOXYGENASE-3
NHEJ: non-homologous end joining recombination
Ni-NTA: nickel –NTA matrix
NLS: nuclear localization signal
No-0: Nossen-0 ecotype
NOR: Nucleolar organizer region

O
ORC1: ORIGIN RECOGNITION COMPLEX-1
ORIs: origins of replication
ORF: open reading frame

P
PARP1: POLY (ADP-RIBOSE) POLYMERASE-1
PCR: polymerase chain reaction
PD: proliferation domain
PIN: PIN-FORMED
PIP2;1: PLASMA MEMBRANE INTRINSIC PROTEIN-2,1
PLT: PLETHORA
Pol I: RNA polymerase I
PRC2: Polycomb repressive complex 2
Pre-RC: pre-replication complexes
PTM: Post-translational modifications

Q
QC: quiescent center
qPCR: quantitative PCR

R
RAD51: RADIATION SENSITIVE 51
RAM: root apical meristem
RBR: RETINOBLASTOMA-RELATED
RdDM: RNA-direct DNA methylation
rDNA: ribosomal DNA
RNA: Ribonucleic acid
rRNA: ribosomal RNA
RT-PCR: reverse-transcription PCR

S
SAM: shoot apical meristem
SAM: S-(5-adenosyl)-L-methionine
SCN: stem cell niche
SCR: SCARECROW
SD: short day
SDG7: SET domain group 7
SDG24: SET domain group 24
SET: SU(VAR)3-9, E(z), and TRX domain
Sf-2: San Feliu-2 ecotype.
SHR: SHORTROOT
siRNA: small interfering RNA
sqPCR: semi-quantitative PCR
SOC1: SUPPRESSOR OF CONSTANS OVEREXPRESSION-1
S-phase: DNA synthesis-phase
STM: SHOOT MERISTEMLESS
SU(VAR)3-9: SUPPRESSOR OF VARIEGATION 3-9
SUVH: SU(VAR)3-9 HOMOLOG
Swi2/Snf2: SWITCH/SUCROSE NON FERMENTABLE
SWN: SWINGER
SWP1: SWIRM DOMAIN PAO PROTEIN

T
TAIR: The Arabidopsis Information Resource
tasiRNA: trans-acting siRNA
TD: transition domain
T-DNA: transfer DNA mediated by Agrobacterium
TE: transposable elements
TEL2N: telomere sequences from chromosome 2
TF: transcription factor
TLP: TOPLESS
TPR: TPL-RELATED
TRF: restriction fragment analysis
TSS: transcription start site
TXR: TRITHORAX

U
UCL1: UPWARD CURLY LEAF-1
USCR: unequal sister chromatid recombination

V
VAR1: 45S gene variant 1
VIN3: VERNALIZATION INSENSITIVE-3
VRN2: VERNALIZATION-2

W
WOX5: WUSCHEL-RELATED HOMEOBOX-5
Ws-0: Wassilewskija-0 ecotype
Wu-0: Wü-0 ecotype

X
xg: times gravity

Z
Zu-0: Zurich-0 ecotype

Abstract
Presentación

ABSTRACT

The *Arabidopsis thaliana* genome contains more than 30 genes encoding SET-domain proteins that are involved in methylating various residues of histones. The SET Domain Group (SDG) gene family can be divided into 5 major subgroups based on the homology with Drosophila SET proteins. Here, we analyzed two *A. thaliana* ASH1 homologs, SDG7 and SDG24, which contain a SET domain accompanied by AWS and post-SET domains. Expression analyses demonstrated that SDG7 is expressed in proliferating cells of above and underground tissues, and localized within the nucleus and cytoplasm of root apical meristem (RAM) cells. The *SDG24* gene revealed the presence of several splicing variants. Protein characterization of two SDG24 isoforms, SDG24.1 and SDG24.2, showed differential expression patterns in the root and gametophyte organs. *In vitro* assays showed that recombinant SDG7 and SDG24.1 bind to modified histone residues, but histone methyltransferase (HMT) activity was not detected. Specific nucleolar expression of SDG24.1 was identified in RAM cells. Analysis of SDG7 and SDG24 null mutants revealed alterations in the cell division and the DNA replication pattern in the RAM, as well as major alterations of ribosomal DNA (rDNA) repeats contained in young and differentiated tissues of the sporophyte. The *sdg7-5* mutant presented an increase of active rDNA copies associated with the nucleolar-organizing region of chromosome 4 (NOR4), whereas the *sdg24-2* mutant displayed a loss of inactive rDNA copies associated with the NOR2 (chromosome 2). The loss of NOR2-inactive variants of *sdg24-2* is epistatic to *sdg7-5* and to *atxr5/6* mutants (H3K27me1 HMTases) regarding the rDNA phenotype. We therefore propose that SDG7 and SDG24 have a major function in the control of NORs stability by regulating rDNA copy number.

PRESENTACIÓN

El genoma de *Arabidopsis thaliana* contiene más de 30 genes que codifican proteínas con un dominio SET, de las cuales varias han sido asociadas a la metilación de diversos residuos de histonas. La familia génica que agrupa el dominio SET (SDG) puede ser dividida en 5 grandes subgrupos basados en la homología de secuencias con proteínas SET de Drosophila. En este trabajo se han analizado dos homólogos de ASH1 en *A. thaliana,* SDG7 y SDG24, los cuales poseen un dominio SET acompañado de un pre- y post dominios SET. Análisis de expresión han demostrado que SDG7 se expresa en células proliferativas de órganos bajo tierra y aéreos, colocalizando en el núcleo y el citoplasma de células del meristemo apical de la raíz (RAM). El estudio del gen *SDG24* reveló la presencia de varias variantes de *splicing*. La caracterización de dos isoformas de la proteína SDG24, SDG24.1 y SDG24.2, demostró un patrón de expresión diferencial en la raíz y órganos del gametofito. Ensayos *in vitro* con las proteínas recombinantes SDG7 y SDG24 revelaron su unión a varios residuos de histonas modificados, aunque la actividad metiltransferasa de histonas (HMT) no ha sido detectada. Adicionalmente, SDG24.1 presentó un patrón específico de expresión en el nucléolo de células del RAM. El análisis de mutantes nulos de expresión de SDG7 y SDG24 reveló la presencia de alteraciones del patrón de división celular y de replicación del DNA en células del RAM, como también alteraciones importantes en las repeticiones de DNA ribosomal (rDNA) presentes en tejidos jóvenes y diferenciados del esporofito. Los mutantes de *sdg7-5* presentaron un incremento de las copias activas de rDNA asociadas a la región organizadora del nucléolo del cromosoma 4 (NOR4), mientras que el mutante *sdg24-2* presentó una pérdida de variantes inactivas asociadas al NOR2 (cromosoma 2). La pérdida de las variantes rDNA-NOR2 del mutante *sdg24-2* demostró ser epistática a los patrones de rDNA que presentaron las mutaciones de *sdg7-5* y *atxr5/6* (HMTasas específicas de H3K27me1). Por lo tanto, se propone que tanto SDG7 como SDG24 poseen una importante función en el control de la estabilidad de los NORs mediante la regulación de las copias de rDNA.

Introduction

1.1. Organization of eukaryotic genome

Within the eukaryotic cell, the nucleus is a highly specialized organelle that houses the DNA. The most important function of DNA is to carry genes, the information that specifies all the proteins that make up an organism, including information about when and in what cells are to be made (Alberts *et al.*, 2014). In eukaryotes, the genome is made of enormously long linear DNA molecules that need to be packaged in order to fit inside the cell nucleus. To do so, they associate with proteins and form the chromatin. The nucleosome is the structural repeating unit of chromatin, which consists of 147 base pairs of DNA wrapped around an octamer of the four core histone proteins — H2A, H2B, H3, and H4. A fifth histone class, the liker histone H1, associates with DNA between single nucleosomes, establishing a higher level of organization (**fig. 1.1**). This 'beads-on-a-string' organization forms the classical 30 nm strand (Olins *et al.*, 1975; Maeshima *et al.*, 2014) though a recent study reveals that chromatin is a disordered 5-to-24-nanometer-diameter chain whit many different particle arrangements, that bends at various lengths to achieve structural compaction (Ou *et al.*, 2017).

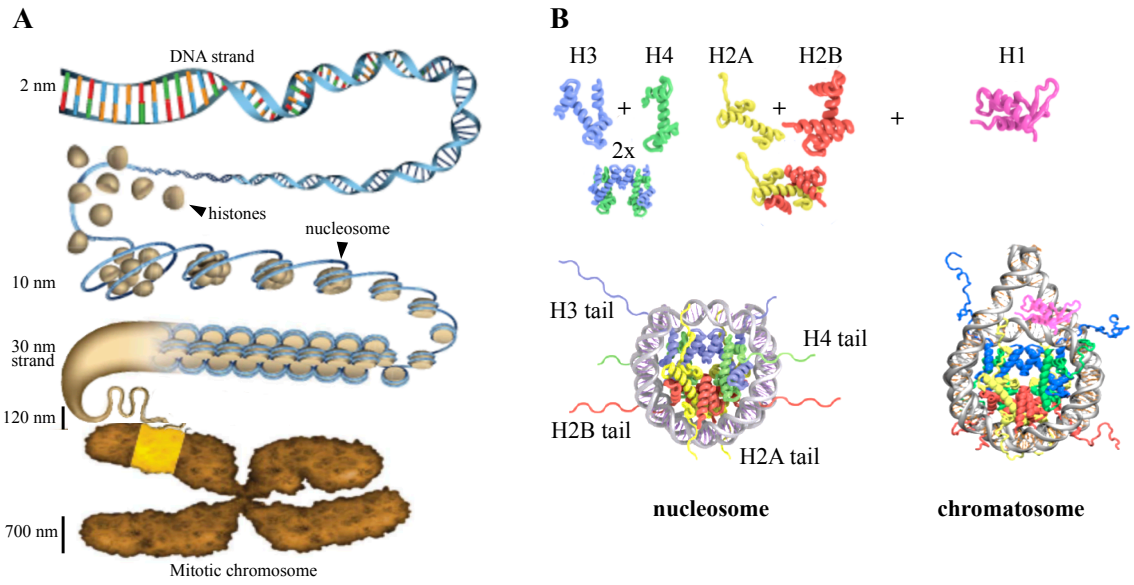


Fig 1.1. Different levels of chromatin compaction. (**A**) Hierarchical chromatin folding model from mitotic chromosome to naked DNA. The 2-nm long DNA molecule is folded around the nucleosome (10-nm), and forms a 30- nm "beads-on-a-string strand. Higher order of compaction keeps folding the DNA into 120-nm chromonema, to 300-nm chromatid (not shown) and 700-nm helical loops in mitotic chromosomes. (**B**) The nucleosome core particle consists of 147 bp of DNA wrapped around an octamer of histones (H3, H4, H2A and H2B, 2 copies of each) with their N-terminal tails protruding from the globular domains. When linker histone H1 is added, keeps in place the DNA around nucleosome (chromatosome), and helps to package the chromatin into higher order structures. Adapted from Bowman and Poirier, 2015; Draizen *et al.*, 2016; Ou *et al.*, 2017.

Local modifications of the nucleosome structure and organization of chromatin fibers translate into the formation of two major chromatin domains in interphase nuclei that can be microscopically detected, e.g. by DAPI staining: bright regions called heterochromatin and weakly stained regions called euchromatin. Euchromatin is gene-rich and decondensed during interphase, whereas heterochromatin is gene-poor and rich in repetitive sequences, and remains mostly condensed throughout the cell cycle (Hsieh and Gage, 2005). Different chromatin compaction levels raise challenges for DNA processes, such as replication, transcription, recombination and repair. Hence, the DNA in order to be accessible, the chromatin needs to be remodeled, and this can be achieved in different ways. For instance, positioning of nucleosomes on DNA can be disrupted and reconfigured to 'open' the chromatin by ATP-dependent remodeling complexes including DECREASED DNA METHYLATION-1 (DDM1) and the helicase Swi2/Snf2 (Tang *et al.*, 2010; Zemach *et al.*, 2013). Also, the histone composition of nucleosomes can be altered by replacement of canonical histones with histone variants, such as H3.1 for H3.3, or by adding a variety of post-translational modifications (PTMs) into the histone amino-terminal tails (Fischle *et al.*, 2003; Gurard-Levin and Almouzni, 2014; Rivera *et al.*, 2014; Kawashima and Berger, 2014).

PTMs are covalent alterations on histones that can be dynamically added and removed enzymatically, and includes phosphorylation (p), methylation (me), acetylation (ac), and ubiquitylation (ub), among others (Bannister and Kouzarides, 2011). These marks, have the potential to alter histone/DNA and histone/histone interactions, and thus provide a means for transiently targeting changes in nucleosome dynamics and gene expression (Bowman and Poirier, 2015). For instance, active transcription has been associated to lysine (K) acetylation by changing electrostatic interactions between DNA and histones (Bannister and Kouzarides, 2011; Bowman and Poirier, 2015), while the methylation on lysine and arginine (R) residues are not likely to affect chromatin structure, rather to create binding sites for regulatory proteins (Gurard-Levin and Almouzni, 2014). In plants, histone lysine methylation can occurs at several residues on histone H3 (K4, K9, K27, K36) and histone H4 (K20) (**Fig 1.2**). All these lysines can be mono (me1), di (me2) or tri-methylated (me3), and the presence of different methylation states can lead to different chromatin signatures (Kouzarides, 2007; Thorstensen *et al.*, 2011).

In Arabidopsis, methylation on H3K9me2 and H3K27me1 are generally associated with heterochromatin, and the appearance of H3K4me2/me3, H3K36me3 and H3K9me3 marks with euchromatic regions (Roudier *et al.*, 2011; Zhang and Ma, 2012; Desvoyes *et al.*, 2014) Nonetheless, individual histone marks are not static indicators of transcription or repression. A recent study in Arabidopsis reported a combinatorial analysis including histone variants,

DNA methylation and histones PTMs defining nine different chromatin states far from the classical euchromatin and heterochromatin conception (Sequeira-Mendes *et al.*, 2014). In this study, each state was associated with distinct levels of DNaseI chromatin accessibility and with described functional genomic elements including promoters, gene bodies of active and repressed genes, and two classes of heterochromatin. Interestingly, active genes include enrichments in H3K4me2/me3 and several acetylations in histones H3 and H4, and repressed genes showed high enrichment of H3K27me3 combined with H3K4me2/me3 marks in a H3.1-rich nucleosome context.

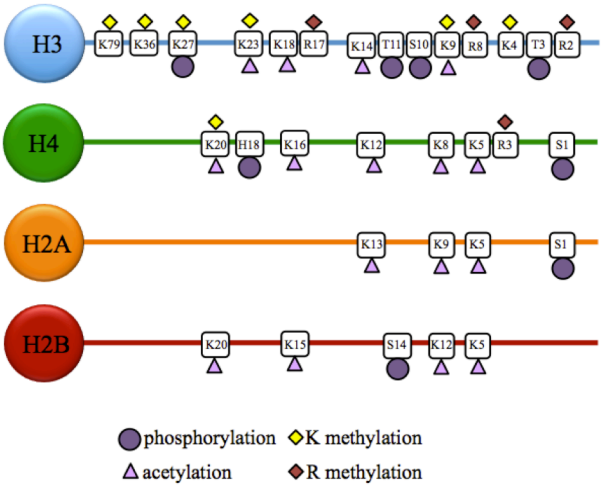


Fig 1.2. Histone tail modifications. Histones carry diverse covalent post-translational modifications (PTMs) at their N-terminal tails such as acetylation, phosphorylation and methylation. Acetylation and phosphorylation can affect DNA– protein interaction, while methylation does not alter the charge of the modified residues on histone tails. In Arabidopsis, Lysine (K) methylation have been described in histone H3 (K4, K9, K27, and recently at K23) and also at histone H4 (K20), though the presence of any form of H4K20 methylation has been questioned. Other common histone PMTs are represented as well. R: arginine; S: serine; T: threonine; H: histidine. Adapted from Liu et al., 2010; Trejo-Arellano et al., 2017.

Deposition and removal of histone PTMs is performed by histone modifying enzymes such as histone acetyltransferases (HATs), histone methyltransferases (HMTs) and kinases (known as the "writers") or histone deacetylases (HDACs), histone demethylases (HDMs) and phosphatases (known as the "erasers") (Onder *et al.*, 2015). In particular, lysine HMTases are proteins with a conserved SET domain responsible for the enzymatic activity. The SET domain is a 130-140 amino acid, evolutionary well-conserved sequence motif that was initially characterized in Drosophila proteins SU(VAR)3-9, ENHANCER OF ZESTE (E(z)), and TRITHORAX (TRX). In Arabidopsis, the SET Domain Group (SDG) protein family have been classified into 5 groups base on their sequence similarity with Drosophila SET proteins: **I,** E(z) homologs; **II**, ABSENT, SMALL, OR HOMEOTIC DISC-1 (ASH1) homologs and related; **III**, TRX homologs and related; **IV**, Arabidopsis TXR related 5 (ATXR5) and ATXR6 homologs; and **V**, SU(VAR)3-9 homologs and related (Baumbusch *et al.*, 2001; Thorstensen *et al.*, 2011). Major activities and interactions for SDG proteins are summarized in **Table 1.1**.

Table 1.1. Properties of Arabidopsis SET domain proteins

Class	Name	SDG number	AtG ID	Histone mark	Method of HMT act.	Interactions/ Activity	References
I- E(Z)	MEA	SDG5	AT1G02580	H3K27me3	ChiP	MEA/SWN; MEA/FIE PHE1 DME and MET1 FIS2 PcG	Spillane et al., 2000; Kohler et al., 2005; Makarevich et al., 2006; Wang et al., 2006; Hennig and Derkacheva 2009
	CLF	SDG1	AT2G23380	H3K27me3	ChiP	EMF2/CLF AG and STM locus ATX1/CLF UCL1/CLF EMF2 PcG and VRN2 PcG complex	Chanvivattana et al., 2004; Makarevich et al., 2006; Schubert et al., 2006 Saleh et al., 2007; Jeong et al., 2011 Hennig and Derkacheva 2009
	SWN	SDG10	AT4G02020	H3K27me3	ChiP	MEA/SWN EMF2 PcG VRN2 PcG	Makarevich et al., 2006; Wang et al., 2006 Hennig and Derkacheva 2009
II- ASH1	ASHH1	SDG26	AT1G76710	H3; H4 H3K4me3 H3K36me1	IVo, PE ChIP	ASHH1/ASHH2 ASHH1/ATX1 ASHH1/ EDA3 ASHH1/ HSP40/DnaJ SOC1 locus	Zhao et al., 2006; Xu et al., 2008; Valencia-Morales et al., 2012; Berr et al., 2015; Liu et al., 2016
	ASHH2/ ESF	SDG8	AT1G77300	H3; H4; H3K36me2/me3; H3K4me3 H3K4me1 reader	IVo; PE; ChIP	ASHH1/ASHH2 FLC Gametophyte regulation Shoot branching	Xu et al., 2008; Dong et al., 2008 Grini et al., 2009; Ko et al., 2010; Valencia-Morales et al., 2012; Hoppman et al., 2011; Cazzonelli et al., 2014; Berr et al., 2015
	ASHR3	SDG4	AT4G30860	H3K36me2	-IVm/h; ChIP	AMS (TF) in Sporophyte dev. Target of E2F	Thorstensen et al., 2008; Kumpf et al., 2014
	ASHH3	SDG7	AT2G44150		-IVo/m/h; -PE	K3me1/me2 MTase of AtPIP2,1 (aquaporin) VIN3 and lncRNAs COLDAIR- COOLAIR	Sahr et al., 2010; Lee et al., 2015
	ASHH4	SDG24	AT3G59960	-	-		
III- TRX	ATX1	SDG27	AT2G31650	H3K4me3	IVp; ChIP	ASHH1/ATX1; ATX1/CLF ATX1/ATXR7 for FLC Stress response NCED3 RNA Pol II recruitment	Saleh 2007: Pien et al., 2008; Tamada et al., 2009; Ding et al., 2011a; Ding et al., 2011b Valencia-Morales et al., 2012
	ATX2	SDG30	AT1G05830	H3K4me2	ChIP	FLC expression	Pien et al., 2008; Saleh et al., 2008; Shafiq et al., 2014
	ATX3	SDG14	AT3G61740	H3K4me2/3	PE; ChIP	ATX3/4/5	Chen et al., 2017
	ATX4	SDG16	AT4G27910	H3K4me2/3	PE; ChIP	ATX3/4/5	Chen et al., 2017
	ATX5	SDG29	AT5G53430	H3K4me2/3	PE; ChIP	ATX3/4/5	Chen et al., 2017
	ATXR3	SDG2	AT4G15180	H3K4me3	PE; ChIP; IVh/p	FLC expression Root branching	Guo et al., 2010; Berr et al., 2010; Yun et al., 2012; Yao et al., 2013
	ATXR7	SDG25	AT5G42400	H3K4me1/2/3 H3K36me2	ChIP; IVo	ATX1/ATXR7 in regulation of FLC expression	Tamada et al., 2009; Berr et al., 2009
IV-	ATXR5	SDG15	AT5G09790	H3K27me1	IVh/p; CH	ATXR5/6	Jacob et al., 2009
	ATXR6	SDG34	AT5G24330	H3K27me1	IVh/p; CH	ATXR5/6; Sporophyte dev.	Jacob et al., 2009; Raynaud et al., 2006
V- SU(VAR)3-9	SUVH1	SDG32	AT5G04940	H3K9me2 H3K4me3	CH, -PE, -IVh	Expression by CHH-DNA binding	Naumann et al., 2005; Ebbs and Bender 2006; Li et al., 2016
	SUVH2	SDG3	AT2G33290	H3K9me1/2* and H3K27me1*, H3K27me2⁑, H4K20me1*	PE, CH, IVo*	SUVH2/SUVH9 Pol V occupancy trough RdDM binding	Naumann et al., 2005; Ebbs and Bender 2006; Johnson et al., 2008; Liu et al., 2014; Jing et al., 2016
	SUVH3	SDG19	AT1G73100	-	-IVh	TPL/TPR interactor	Ebbs and Bender 2006; Causier et al., 2012
	SUVH4/ KYP	SDG33	AT5G13960	H3; H3K9me1/2	IVh/p; ChIP; CH	SUVH4/5/6 SUVH4/5/6/HDA6 DNA binding at CHGme	Jackson et al., 2002; Naumann et al., 2005; Johnson et al., 2007; Ebbs and Bender 2006; Yu et al., 2017
	SUVH5	SDG9	AT2G35160	H3; H3K9me1/2; H2A	IVh; ChIP	SUVH4/5/6 SUVH4/5/6/HDA6	Jackson et al., 2004; Naumann et al., 2005; Ebbs and Bender 2006; Yu et al., 2017
	SUVH6	SDG23	AT2G22740	H3; H3K9me1/2	IVh/p; ChIP	SUVH4/5/6 SUVH4/5/6/HDA6 DNA binding at CHGme	Jackson et al., 2004; Naumann et al., 2005; Ebbs and Bender 2006; Johnson et al., 2007; Yu et al., 2017
	SUVH7	SDG17	AT1G17770		-IVh		Ebbs and Bender 2006
	SUVH8	SDG21	AT2G24740		-IVh	Target of passenger strand of miR171a	Ebbs and Bender 2006; Manavella et al., 2013
	SUVH9	SDG22	AT4G13460		-IVh/o;	SUVH2/9 Pol V occupancy trough RdDM binding	Ebbs and Bender 2006; Johnson et al., 2008 Liu et al., 2014; Jing et al., 2016
	SUVH10	SDG11	AT2G05900				
	SUVR1	SDG13	AT1G04050		-IVh	SUVR1/SUVR2 Binding to CHR19/27/28. Silencing by RdDM Ubiquitin binding	Thorstensen et al., 2006; Han et al., 2014; Rahman et al., 2014
	SUVR2	SDG18	AT5G43990		-IVh	SUVR1/2	Thorstensen et al., 2006; Han et al., 2014; Rahman et al., 2014
	SUVR3	SDG20	AT3G03750				
	SUVR4	SDG31	AT3G04380	H3K9me2/3	IVh/p; ChIP	Ubiquitin binding H3K9me3 HMT transposon specific	Thorstensen et al., 2006, Veiseth et al., 2011; Rahman et al., 2014
	SUVR5	SDG9	AT2G23740	H3K9me1, H3K27me2 H3K9me2	ChIP	SUVH5/SWP1; SUVH5/LDL1 demethylase FLC repression DNA binding	Krichevsky et al., 2007 Caro et al., 2012

In vivo association to a histone mark was assessed by: ChiP (chromatin immunoprecipitation), CH (cytohistology) and PE (protein extracts) of mutant lines. *In vitro* (IV) Histone Methyltransferase assay (HMT) was performed with different substrates: oligonucleosomes/ mononucleosomes/free histones/ histone peptides (o/m/h/p). A negative signal (-) in front of IV means negative tested results. (*) Detected by Naumann et al., 2005, was not reproduced by Johnson et al 2008. (⁑) At heterochromatin by Naumann et al., 2005, but detected at euchromatin by Johnson et al 2008. MTase: non-histone methyltransferase

In general, the E(z), ATXR5/6 and SUVH proteins play a role in repressing gene/transposon expression through accumulating H3K27 or H3K9 methylation modifications, while the ASH1 and TRX proteins methylate H3K36 and H3K4, thereby activating gene expression (Thorstensen *et al.*, 2011). In particular, the activity of SDG proteins has been described to control different plant processes. For example, the E(Z) protein CURLYLEAF (CLF), four TRX proteins (ATX1, ATX2, ATXR3 and ATXR7), and two ASH1 proteins (ASH1-HOMOLOG-1 (ASHH1) and ASHH2) all act in regulation of flowering time through controlling the expression of FLOWERING LOCUS C (FLC) (Shafiq *et al.*, 2014; Zhao *et al.*, 2015). SDGs are also required in shaping other aspects of plant development such as sporophyte development (Raynaud *et al.*, 2006; Thorstensen *et al.*, 2008; Grini *et al.*, 2009; Berr *et al.*, 2010), shoot branching, leaf size, root length, number of lateral roots (Dong *et al.*, 2008; Yao *et al.*, 2013), and moreover, chromatin signatures given by some SDG HMTases have been implicated in the regulation of ribosomal DNA (rDNA) gene expression inside de nucleolus (Pontvianne *et al.*, 2012).

1.2. The nucleolus

The nucleolus is the most prominent nuclear body on interphase nuclei. When stained with fluorescent DNA dyes, is seen as a dark region within the more brightly stained nuclear chromatin (Lo *et al.*, 2006). For decades, the main purpose of the nucleolus has been attributed to ribosome biogenesis; although increasing data support new functions, including signal recognition, particle assembly, small RNA modification, telomerase maturation, cell stress sensor, and cell cycle and aging control (Boulon *et al.*, 2010; Olson, 2011; Stepinski, 2014; Wang *et al.*, 2016; Pontvianne *et al.*, 2016; van Sluis and McStay, 2017).

The number and shape of nucleoli present within a cell may vary between species. For instance, in mammals several nucleoli with an irregular contour can be found, while in plants is often observed as one nearly spherical entity (**Fig. 1.3**) (Shaw and Brown, 2012; Stepinski, 2014). On the basis of their appearance by electron microscopy (EM), three main nucleolar components (compartments) can be discerned: the fibrillar centers (FCs), the dense fibrillar component (DFC) and the granular component (GC). The FCs are clear areas, partly or entirely surrounded by the DFC; and both are embedded within the GC (**Fig 1.3**). (Shaw and Brown, 2012). This organization may vary as well according to the cell type, cell cycle, physiological state of the cell, transcriptional activity of the nucleolus, impact of biotic and abiotic factors, and to a certain extent according to species (Hernandez-Verdun, 2011b; Stepinski, 2014).

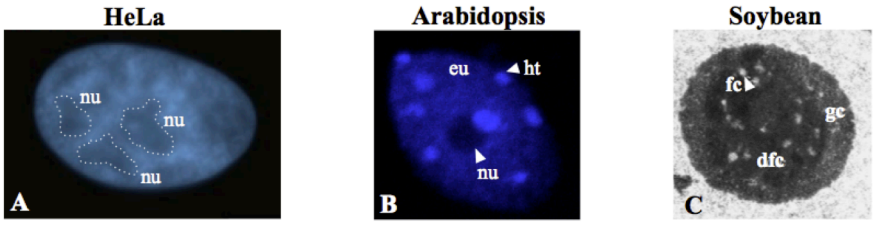


Figure 1.3. The Nucleolus organization. Nucleoli can be detected inside the nucleus as dark regions poorly stained with the fluorescent dye DAPI. (**A**) HeLa nuclei. (**B**) *Arabidopsis thaliana* nuclei. (**C**) The sub-nucleolar compartments from Soybean nucleolus can be detected by electron microscopy (EM). eu: euchromatin; ht: heterochromatin; nu: nucleolus; fc: fibrillar center; dfc: dense fibrillar component; gc: granular component. Images from Hernandez-Verdun, 2011b; Pontvianne *et al*., 2016; Stepinski 2010.

During the cell cycle, the size of nucleolus is in agreement whit the cell's ribosome needs, which is sustained throughout G1/S/G2, been at maximal production in G2. After that, the nucleolus disassembles at the beginning of mitosis and reassembles once again at the exit of mitosis in telophase (Hernandez-Verdun, 2011a). The position of the nucleolus in the nucleus is not random, as it builds around specific chromosomal features, termed nucleolar organizer regions (NORs), which contain the repeated arrays of ribosomal DNA (rDNA). NORs are easily identified in metaphase chromosomes as secondary constrictions, where the chromatin is less condensed (McClintock, 1934). The number of NOR-bearing chromosomes varies in different species, from one chromosome to several chromosome pairs. In Arabidopsis, NORs are the most distal genetic loci of the short arms of chromosomes two and four (NOR2 and NOR4, respectively), where the terminal rDNA genes are capped directly by telomere repeats (Copenhaver and Pikaard, 1996; Stepinski, 2014).

Ribosomes are composed of four ribosomal RNA (rRNA) species: 18S, 5.8S, 25S and 5S. In Arabidopsis, the first three species, encoded by a 45S rDNA unit, are tandemly arrayed at NOR2 and NOR4 and transcribed by RNA polymerase I (Pol I) inside the nucleolus (**Fig. 1.4**). On the other hand, 5S rDNA arrays are dispersed across chromosomes III, IV and V, and opposite to the 45S unit, 5S genes are transcribed at the extranucleolar nucleoplasm by RNA polymerase III (Pol III)(Benoit *et al.*, 2013; Stepinski, 2014). The rDNA represents one of the most abundant DNA repeats, comprising ~5% of Arabidopsis genome (~16 Mb). Each NOR consist of ~700–800 of 45S rDNA copies per diploid cell, and each rDNA gene spans ~10–10.5 kb (Copenhaver and Pikaard, 1996).

In the Arabidopsis Col-0 ecotype, four main 45S rDNA gene types (or variants) have been identified so far based on differences within a repetitive region of the external transcribed spacer located just at the 3′ of 25S rRNA sequences (**Fig. 1.4**) (Chandrasekhara *et al.*, 2016). These four gene types are revealed by PCR amplification of genomic DNA with

specific primers for the 3′ETS region. Three of the rDNA types are abundant (Variant 1 (VAR1), VAR2, VAR3), and one is relatively rare (VAR4), although additional changes can also be found within each variant, adding more complexity to rDNA alleles (Chandrasekhara *et al.*, 2016; Havlova *et al.*, 2016; Rabanal *et al.*, 2017). All four classes of 45S rDNA genes are expressed (rRNA) in newly germinated seeds, but, by 10–14 days after germination and throughout the remainder of vegetative development, the VAR1 class, accounting for ~50% of the total rDNA gene pool, is selectively silenced (Pontvianne *et al.*, 2010).

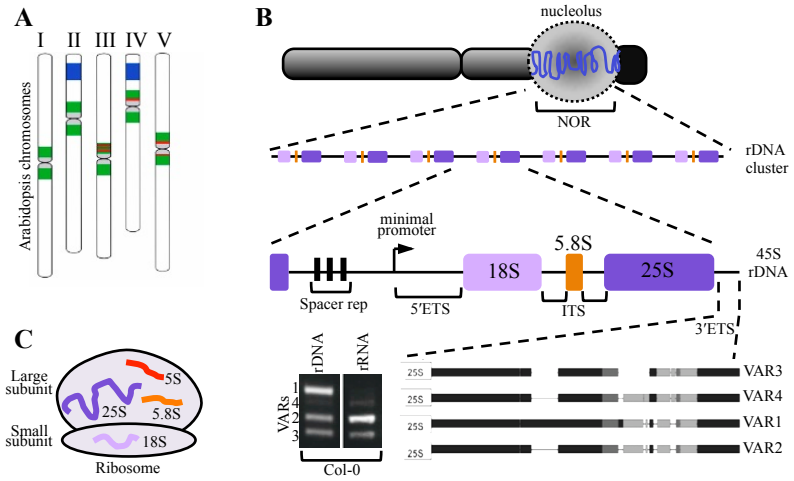


Figure 1.4. Relationships between the nucleolus, NORs, and 45S rRNA gene repeats. (**A**) Schematic representation of the five *Arabidopsis thaliana* chromosomes (2n = 10) in the Columbia accession (Col-0). Chromosome II and IV carry the 45S rDNA loci (Nucleolus Organizer Regions, NOR, blue). The 5S loci (red) are dispersed on chromosomes III, IV and V, in close proximity to centromeric repeats (gray) and inside the pericentromeric domains (green). (**B**) A metaphase chromosome displaying a NOR with the secondary constriction (blue line) cap by telomeric repeats (black). The 45S rDNA genes (18S, 5.8S and 25S) are tandemly arrayed at the NOR clusters. In *A. thaliana*, insertions/deletions in the 3′ external transcribed region (3′ ETS) define 4 major rDNA gene variant types (VAR1-4) observed by PCR amplification. Several days after germination in the Col-0 ecotype, VAR1 expression is silenced (rRNA) and VAR2 increased. (**C**) Drawing of a ribosome displaying the localization of rRNAs distributed along the large (25S, 5.8, 5S) and small (18S) subunits. Modified from Benoit et al., 2013; Pontvianne et al., 2013.

It has been proposed that this selective rDNA silencing (gene dosage), is not regulated by gene-based mechanisms, but rather it depends on a multimegabase scale inactivation of NORs (Preuss *et al.*, 2008; Chandrasekhara *et al.*, 2016; Mohannath *et al.*, 2016). This hypothesis is supported by observations of NORs organization at the chromosomal level. In particular, in Arabidopsis Col-0, all silenced rDNA gene subtypes (VAR1 and a subset of VAR3) mapped to the NOR2, and all active rDNA subtypes (VAR2, VAR4 and the majority of VAR3 genes) localized at NOR4 (Chandrasekhara *et al.*, 2016). Moreover, active rDNA genes are present at, or close to, nucleoli when active, and excluded from it when silenced

(Pontvianne *et al.*, 2013; Chandrasekhara *et al.*, 2016). While NOR4 activation/NOR2 inactivation pattern occurs in Col-0, this is not a general feature across Arabidopsis ecotypes. In the study carried out by Rabanal et al. (2017), they showed that some accessions (Sf-2, Bur-0, Edi-0, Ws-0, and Wu-0) appear to behave like Col-0; other accessions **(**No-0, Ct-1, Can-0, and Hi-0) show the opposite pattern, where NOR2 is exclusively expressed, and two ecotypes **(**Ler-0 and Zu-0) express both NOR2 and NOR4. Furthermore, the observed relative size of the two NOR clusters varies greatly among accessions, although gene copy number variation does not alter rRNA transcription (Rabanal *et al.*, 2017). Different crossed combination between Arabidopsis accessions sharing the same or different NOR activities, correlate with a hierarchical dominance-silencing phenomenon known as nucleolar dominance (Chandrasekhara *et al.*, 2016; Rabanal *et al.*, 2017).

Nucleolar dominance is a common phenomenon in cells of interspecific hybrids of plants and animals that describes the failure to form nucleoli inherited from one parent (Pikaard, 2003). In the same year that secondary constrictions were described as the active NORs, observation in hybrids of the plant genus Crepis showed that only NORs derived from one species would form secondary constrictions (Navashin, 1934). Regardless of whether which species served as the maternal or paternal parent, NORs derived from a number of species could be organized into a dominance hierarchy. Interestingly, secondary constrictions could form once again on both diploid copies of the chromosome at the F2, showing that the affected loci had not been lost or altered in the hybrid. These observations established nucleolar dominance as an epigenetic phenomenon (Pikaard, 2003).

DNA methylation plays a central role in the maintenance of nucleolar dominance. In *Arabidopsis suecica*, the allotetraploid hybrid of *A. thaliana* and *A. arenosa,* silent NORs corresponding to those of *A. thaliana* can be reactivated by 5-aza-2′ deoxycytosine (aza-dC), an inhibitor of cytosine methylation, or by treatment with histone deacetylase inhibitors such as trichostatin A (Chen and Pikaard, 1997; Lawrence *et al.*, 2004; Pontes *et al.*, 2007). Analysis derived from the used of both chemicals resulted in a model describing how cytosine methylation and histone deacetylation specify one another in a self-reinforcing cycle that maintains rRNA gene silencing (Lawrence *et al.*, 2004). Over the years, several proteins involved in maintenance of nucleolar dominance in *A. suecica* have been identified, including HISTONE DEACETYLACE 6 (HDA6), METHYLCYTOSINE-BINDING DOMAIN protein 6 (MBD6), MBD10, the *de novo* DNA methyltransferase DRM2, and proteins of the siRNA-directed DNA methylation pathway (Preuss *et al.*, 2008).

On the other hand, a nucleolar dominance-like phenomenon also occurs in non-hybrid *A. thaliana*, in which distinct rRNA gene variants are selectively inactivated during early

vegetative development (gene dosage compensation, described above). Alteration of silencing patterns has been attributed likewise to the action of HDA6, in a common mechanism observed for *A. suecica,* but also to the action of different chromatin modifiers, including 2 subunits of chromatin assembly factor 1 CAF1 (FASCIATA-1 (FAS1) and FAS2), nucleolin, several H3K9 HMTases (SUVR4, SUVH5 and SUVH6), and H3K27me1 HMTases (ATXR5/ATXR6), all of them altering the chromatin compaction of NOR clusters near the nucleolus (Earley *et al.*, 2010; Pontvianne *et al.*, 2010; Pontvianne *et al.*, 2012; Pontvianne *et al.*, 2013). In general, the mechanisms by which one parental set of NOR-bearing rRNA genes are chosen are still unknown. However, it is clear that a partnership between DNA methylation, histone modifications, and even siRNA pathways determines rDNA gene silencing.

1.3. The cell cycle

The eukaryotic cell cycle is a highly regulated process with the purpose of giving rise to two daughter cells. It is typically divided into four major phases: the post- mitotic gap phase (G1), the DNA synthesis phase (S), the post-synthetic gap phase (G2), and the mitosis (M) (**Fig. 1.5**) (Gutierrez, 2009). In particular at G1, origins of replication (ORIs) are licensed and pre-replication complexes (Pre-RC) are loaded at potential ORIs that will fire in S-phase (Sanchez *et al.*, 2012). At this stage, a first transcriptional wave occurs depending on the activity of the RETINOBLASTOMA-RELATED (RBR) protein and various RBR-interacting E2F proteins (Desvoyes *et al.*, 2006; Ramirez-Parra and Gutierrez, 2007). The presence of RBR induces the up-regulation of genes encoding, among others, all pre-RC factors (CDC6, CDT1, MCM2-7, and ORC1-6, excepting ORC5) and favors recruitment of histone HDACases, HMTases and DNA methyltransferases (Desvoyes *et al.*, 2014).

Initiation of genome replication marks the beginning of S-phase, and lasts until the entire DNA and chromatin components are duplicated. This includes several processes where chromatin becomes accessible, implying nucleosome remodeling, changes in specific histone modifications, and the participation of histone chaperones (MacAlpine and Almouzni, 2013; Desvoyes *et al.*, 2014). G2 is an intermediate phase between S-phase and mitosis, which will ensure that the cell had all DNA and intracellular components properly duplicated prior to enter to mitosis. In plants, G2 progression is given by the activity of several cyclin-dependent kinase/cyclin complexes (CDKA/CYC and CDKB/CYC). Expression of these genes, together with others showing similar G2/M-phase-specific pattern (e.g. R1R2R3-Myb) are regulated trough the presence of mitosis-specific activator (MSA) elements (Ito *et al.*, 1998; Haga *et al.*,

2011). Particularly in Arabidopsis, MYB3R1/4 proteins act as transcriptional activators of G2/M-specific genes, including CYCB1 genes, CDC20.1 and KNOLLE, a gene essential for cell plate formation (Haga *et al.*, 2007; Haga *et al.*, 2011). Following G2, in mitosis the duplicated chromatin becomes condensed; forming visible chromosomes that will be aligned in metaphase and distributed equally to the newborn cells. At the final stage of cell cycle, cytokinesis will procure the separation of the daughter cells by a coordinated action between vesicles and cytoskeletal components in order to form the cell plate that will divide the cell (Gutierrez, 2009).

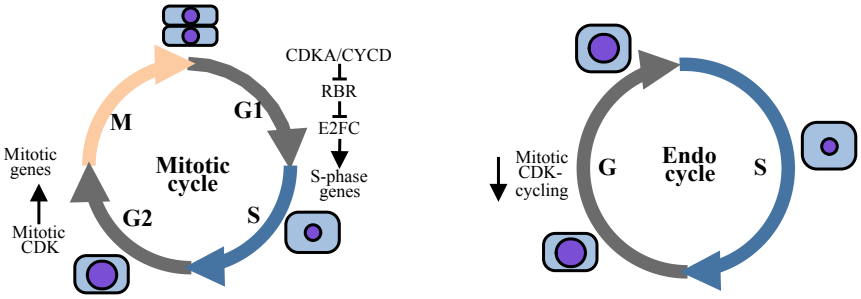


Fig 1.5. The Cell cycle. Representation of the two basic cell cycles. In the mitotic cycle, new-born cells pass through G1, S and G2 before dividing and produce two daughter cells. CDK/cyclin complexes drive the G1/S and G2/M transitions. In the endocycle, the M phase does not occur and cells undergo a doubling of its nuclear DNA content in each endocycle, frequently associated with an increase in cell mass. A decrease in mitotic CDK/cyclins promotes endocycle entry. M: mitosis; S: synthesis; G/G1/G2: Gap phases. Adapted from Gutierrez, 2009.

In addition to the classical cell cycle described above, plant and animal cells also posses a different cell cycle mode called endoreplication. It consists of replication of DNA without any subsequent mitosis and cytokinesis, and thereby the cell becomes polyploid (**Fig. 1.5**) (Gutierrez, 2009). In Arabidopsis, the action of endoreplication is observed in tissues such as hypocotyls (embryonic shoot), roots, leaves (mesophyll, pavement and trichome cells), fruits, and seed endosperm (Edgar *et al.*, 2014). The endocycle uses much of the same machinery that regulates the transition from G1 to S phase in mitotic cell cycles. Mitosis-to-endocycle transition is affected by E2F-RBR complexes, and triggered by downregulation of the mitotic-CDKs permitting a lower, but continuous, activity of CDK acting in G1-to-S progression (S-CDKs). However, S-CDKs are periodically inactivated to allow a transition to G1-like gap phase, where pre-replicative complexes are loaded (Edgar *et al.*, 2014).

1.4. Plant organogenesis: the Arabidopsis model

Plants have an extraordinary developmental plasticity as they continuously form organs in a post-embryogenic manner (Perianez-Rodriguez *et al.*, 2014). Aboveground and belowground organs are formed from two distinct stem cell populations located at the shoot and root apical meristems (SAM and RAM, respectively)(**Fig. 1.6**). Primary meristems are formed as a result of embryogenesis, and upon activation during germination they start generating the main root, leaves and flowers (Capron *et al.*, 2009; Perianez-Rodriguez *et al.*, 2014). Being sessile, plants have to be especially equipped to translate environmental cues into developmental responses. New growth axes need to be established to in order to be compatible with the environment, and this is achieved by establishing secondary meristems (Agusti and Greb, 2013). Formation of organs will respond to a crosstalk between hormones and regulation of gene expression downstream of epigenetic regulation (Perianez-Rodriguez *et al.*, 2014).

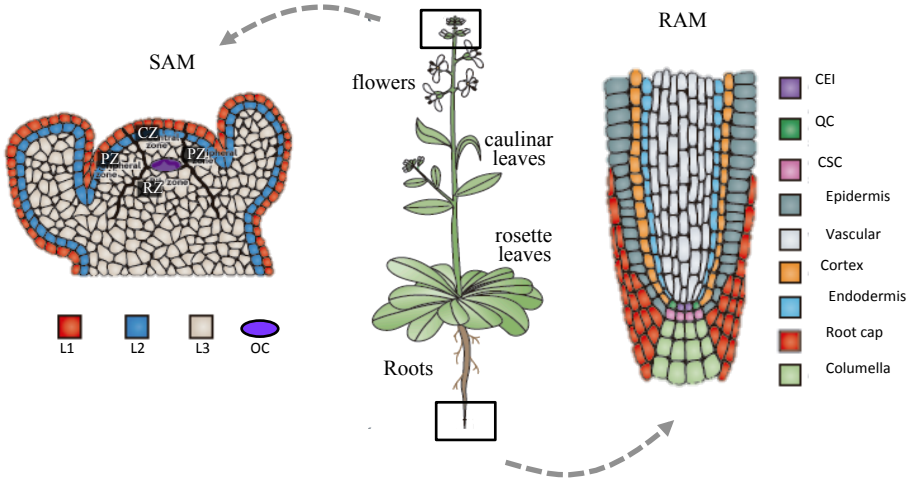


Fig 1.6. Arabidopsis primary meristems. Two meristem populations are formed during embryogenesis, one in the shoot (SAM) and the other in the root (RAM). Both meristems contain niche cells and stem cells to maintain totipotency capacity to produce new organs (e.g leaves, flowers and roots) in a post-embryogenic manner. The SAM contains three layers (L1, L2 and L3) and three developmental zones (peripheral zone, PZ; central zone, CZ; and rib zone, RZ). Cells of the organizing centre (OC) of the SAM are specified at the junction of the three developmental zones and function to maintain stem cells in the shoot. The RAM is radially symmetric and consists of central niche cells (the quiescent centre) surrounded by stem cells initials. In the root, each stem cell population gives rise to one or two cell types. QC: quiescent center; CEI: cortex-endodermis initials; CSC: columella stem cells. Adapted from Sparks *et al*., 2013.

Here is this study we took advantage of *Arabidopsis thaliana* as model to follow plant organogenesis as it can be easily cultivated in confined laboratory conditions and possesses a short life cycle (around two months) comprising the sporophyte (vegetative structures) and the gametophyte (gametes production). Also, Arabidopsis exhibits other advantages for

genetic studies including a full sequenced and well annotated genome (around 125 Mb), is easily transformed, many knockouts mutants are available, and tolerates mutations in key chromatin genes that are lethal in other organisms (Van Norman and Benfey, 2009; Koornneef and Meinke, 2010).

1.4.1. The root

The Arabidopsis root is a thin (~150 μm) and highly organized organ that can be easily grown in non-soil media and imaged by microscopy, characteristics that justify the use of the root as a tool to understand plant organogenesis (Petricka *et al.*, 2012; Geldner and Salt, 2014). Overall, plant survival depends on appropriate root development, growth, and function, as it provides to the plant water and nutrients, and for the necessary anchoring of aerial part to the soil. The root is formed during embryo development and after germination a group of actively proliferating cells in the RAM allows the continuous growth of the organ (Petricka *et al.*, 2012).

Structurally, the RAM is maintained by an organizing center termed the quiescent center (QC), located at the tip of the apical meristem (**Fig. 1.6**). The QC comprises 2-4 cells that rarely divide and is surrounded by the stem cell initials, which together form the stem cell niche (SNC) (Benfey and Schiefelbein, 1994). The stem cells will produce the different root cell types. In the rootward part of the QC, columella stem cells will produce the distal columella cells. The epidermis and lateral root cap, the most external cell layers of the root cylinder, are originated from the epidermis/lateral root cap (LRC) initial cells, while the cortex and endodermis layers will result from an asymmetric division of cortex/endodermis initial daughter cells. Concerning the vascular tissue, vascular initials cells, located in the shootward part of the QC, will generate xylem and pluripotent procambial cells that will later generate the phloem through periclinal divisions (Petricka *et al.*, 2012).

In the SCN, QC cell fate is maintained by the jointly action of the PLETHORA (PLT) (Aida *et al.*, 2004; Galinha *et al.*, 2007); and SCARECROW/ SHORTROOT (SCR/SHR) pathways (Nakajima *et al.*, 2001; Sabatini *et al.*, 2003), while the transcription factor WOX5 is required in the QC to maintain the undifferentiated state of the surrounding stem cells (Galinha *et al.*, 2007; Forzani *et al.*, 2014). Additionally, the ETHYLENE RESPONSE FACTOR-115 (ERF115) also contributes to controls QC division and stem cell replenishment by two antagonistic mechanisms: restriction of QC division trough proteolysis by the APC/$C^{CCS52A2}$ ubiquitin ligase, and QC proliferation by brassinosteroid-dependent ERF115 expression (Heyman *et al.*, 2013). A flowing gradient of auxin is also an important regulator

of QC specification combined with the action of PIN family members. Cytokinins also act to regulate auxin distribution along the root apical meristem by regulating PINs and auxin carriers, resulting in the late regulation of mitotic activity at the QC (Zhang *et al.*, 2013). The hormonal crosstalk between auxin, cytokinin and gibberellin regulates the expression of auxin response repressors as cells transition from the meristem to the elongation zones. Besides to QC maintenance, the SHR:SCR complex is required for endodermal and cortical differentiation. In the presence of auxin, the SHR:SCR complex directly activates CYCD6;1 expression and promotes asymmetric cell division at the initials or its daughter cell (Wilson *et al.*, 2013).

As the root grows, continuous production of the same cell types in approximately the same numbers and in the same places is observed. This results in the different stages of development being present in zones along the root axis (**Fig. 1.7**). These zones and their boundaries can be well defined in several areas based on their characteristic cellular activities: (i) the root apical meristem (RAM) zone, which comprises the proliferation domain and the transition domain, were division capacity slowly decreases as cells are located close to the end of the RAM; (ii) the elongation zone (EZ) in which cells primarily expand; and (iii) the differentiation zone (DZ) in which cells attain their differentiated characteristics and final size (Verbelen *et al.*, 2006; Wilson *et al.*, 2013)

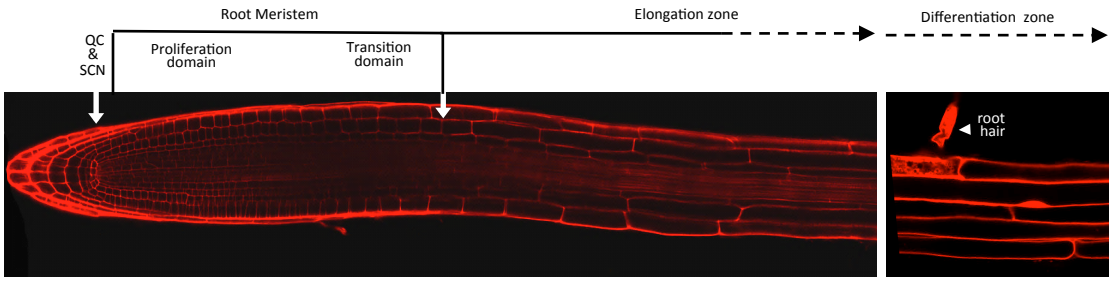


Fig 1.7. Arabidopsis root developmental areas. The root meristem (RAM) is divided in two different domains, the proliferation domain, where cells actively divide, and the transition domain, where mitosis are less frequent. In the elongation zone cells start a rapid elongation and do not divide again. In the differentiation zone, cells start to specialize (e.g. root hairs from the epidermal cell layer). The QC and SCN cells are located in the tip of the RAM.

Once the cells have reached their final size, they acquire maturity characteristics and specialize. A particular case is the change observed at the epidermal cell layer. Upon maturation the epidermis consist of two distinct cell types: root hairs (trichoblast) and no-hair cells (atrichroblasts) (Lofke *et al.*, 2015). Their patterning responds to a position-dependent mechanism relative to the cortex cells regulated by a network of transcription factors, although a cellular dimorphism between trichoblast and atrichroblasts already exists in the

meristematic zone: hair cells are shorter, less vacuolated, and display a higher cell division rate (Berger *et al.*, 1998; Balcerowicz *et al.*, 2015). Another characteristic of differentiation in the root, is cell endoreplication, starting early in root development, precisely at the boundary between meristematic and elongation zone and preceding rapid cell expansion (Hayashi *et al.*, 2013).

Another stage of root growth is root branching, critical to adapting to local soil environments. This relies on the *de novo* formation of lateral roots (LRs), which are originated from a subset of pericycle cells adjacent to the xylem pole "primed" at the transition zone in response to an auxin gradient (Wilson *et al.*, 2013). The resulting asymmetric cell division creates two central daughter cells and larger flanking cells with different cell fates. A positive-feedback loop of auxin reinforces the continuous growth and division throughout the different cell layers (endodermis, cortex and epidermis) creating a dome-shaped LR primordium (LRP) that eventually forms a new meristem. This process is controlled by several auxin receptors, transcription factors and oxygen reactive species (Wilson *et al.*, 2013; Lavenus *et al.*, 2013; Vermeer and Geldner, 2015; Orman-Ligeza *et al.*, 2016; Fernandez-Marcos *et al.*, 2017).

1.4.2. Aerial organs

The organization of the SAM explains how plants are able to grow while still producing aerial organs. The SAM is established in the embryo and has specific structural characteristics (**Fig. 1.6**). In the SAM, cells in the central zone divide to maintain a pluripotent stem cell population, whereas in the peripheral zone, cells are competent to differentiate. During vegetative development, the SAM generates leaf primordia directly from its flanks in a typical spiral pattern with very short internodes between them (Fletcher, 2002; Vanhaeren *et al.*, 2015). The conversion of a leaf primordium into a mature leaf is usually described as consisting of two partially overlapping phases, cell proliferation and cell expansion. In the first phase, proliferation occurs throughout the entire primordium and generates new cells in which size remains relatively constant and small. In the second phase, cell division in the developing leaves has ceased and further growth is mainly achieved by cell expansion, resulting in a large increase in cell size (Gonzalez *et al.*, 2012).

In Arabidopsis cell expansion is frequently associated with endoreduplication (Edgar *et al.*, 2014). Cell proliferation, cell expansion and the coordination between these two processes by complex networks of genes determine the final shape and size of a leaf (Gonzalez *et al.*, 2012; Vanhaeren *et al.*, 2015). In addition to leaf, trichomes (leaf hairs) also follow an

endoreduplication pattern. Trichomes are single-celled hairs that have a stellate appearance with a stalk and three or four branches. Development of this striking morphology starts with the switch of a protodermal cell in the leaf primordium from mitotic to endocycling, during which nuclei eventually reach 32C in strict association with increased branching. Trichomes emerged from both adaxial/abaxial (above and below) side of leaf epidermis, and play a role in protecting the plant against herbivores, from UV irradiation, or reduction of transpiration (Tominaga-Wada *et al.*, 2011). At the end of the vegetative phase, environmental and endogenous factors prompt the plant to undergo the transition to flowering and reproductive development. During this phase the stem elongates and secondary SAMs are formed. At this stage, The SAM is transformed is into the inflorescent meristem (IM) where floral meristems will generate on the flanks of both the primary and secondary SAMs.

1.4.3 Flower and gametophyte development

During the life cycle of Arabidopsis haploid gametophytes form within the diploid sporophytic tissues that constitute the sexual organs of the flowers (Yadegari and Drews, 2004). Initially, flowers develop from the floral meristem (FM), made of small, spherical-shaped mounds of cells, which produce four types of lateral organs in concentric rings called whorls. Sepals are initiated first in the outermost whorl, followed by petals in the second whorl, and stamens in the third whorl. The FM is then consumed in the formation of the central carpel, which form the gynoecium that ultimately encloses the seeds of the next generation (**Fig. 1.8**) (Fletcher, 2002). Flower stereotypic development is essentially controlled by a network of coordinating growth and cell-fate determination, which integrates several hormonal signals, transcriptional regulators, and mechanical constraints in order to specify each whorl and gametophytes (Alvarez-Buylla *et al.*, 2010; Prunet and Jack, 2014; Denay *et al.*, 2017).

In angiosperms, including Arabidopsis, the male gametophyte (the pollen grains) develops within the anthers, and the female gametophyte (megagametophytes or embryo sac) develops within the ovule inside the carpel (**Fig. 1.9**) (Sundaresan and Alandete-Saez, 2010). Male and female gametes will be produced and mature together as the flower develops. Several morphological studies correlated flower stages to particular gametophyte development, although these events are considered approximate (Smyth *et al.*, 1990; Bowman, 1994; Schneitz *et al.*, 1995; Alvarez-Buylla *et al.*, 2010).

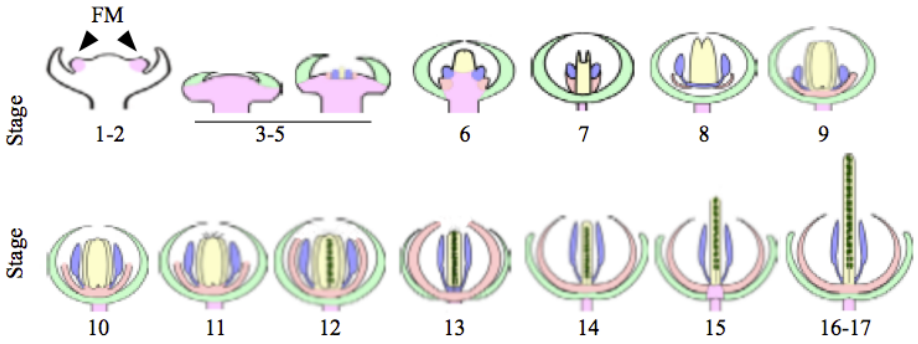


Figure 1.8. Schematic representation of Arabidopsis flower development. During Arabidopsis reproductive phase, flowers originate from the floral meristem (FM) to form the gametes. Flower primordium is established at stages 1 and 2. The sepal primordia become visible at stage 3, the stamen primordia at stage 4, and carpel and petal primordia at stage 5. From stage 6–8, the floral organ primordia enlarge and begin to differentiate. At the stage 9, stigmatic papillae arise at the top of the gynoecium. At stage 12, petals are similar in length to stamens. Megasporogenesis and microsporogenesis begin at stage 7, and the pollen matures shortly before anthesis at stage 13 (gametogenesis is more detailed at fig. 1.9). At stage 14, pollination takes place and flowers opens. Between stages 15 to 17, silique grows and embryos fully develop inside the seed. After siliques desiccation, seeds fall at stage 20 (not shown). In pink: FM; green: sepals; petals: bright pink; stamens: purple; gynoecia: yellow; ovules: dark spheres inside the gynoecium. Adapted from Alvarez- Buylla *et al*., 2010; and Prunet and Jack, 2014.

Regarding male development, mature pollen is produced in the anthers by initial meiosis of pollen mother cells to form tetrads of haploid microspores. Then, microspores enlarge and undergo an asymmetric division forming two cells, in which one is much smaller than the other. The larger cell is called the vegetative cell, and the smaller cell is called the generative cell. The vegetative cell does not divide again but eventually will form the pollen tube. The generative cell is immersed inside the cytoplasm of the vegetative cell and undergoes mitosis once again to form the two sperm cells. The sperm cells are identical male gametes that will fertilize female gametic cells (McCormick, 2004).

In the case of female gametophyte, it is embedded within several layers of female ovular tissues, including the funiculus, the chalaza and the nucellus, the last one being where the megaspore lineage will develop (Schneitz *et al.*, 1995). Inside the ovule, a meiotic division produces four haploid megaspores, one of which develops into the functional megaspore via a cytokinin spatial signal, which after three nuclear divisions generates the seven-celled gametophyte: three antipodal cells, two synergid cells, one egg cell, and one central cell containing two polar nuclei that fuse prior to or during fertilization (Lawit *et al.*, 2013).

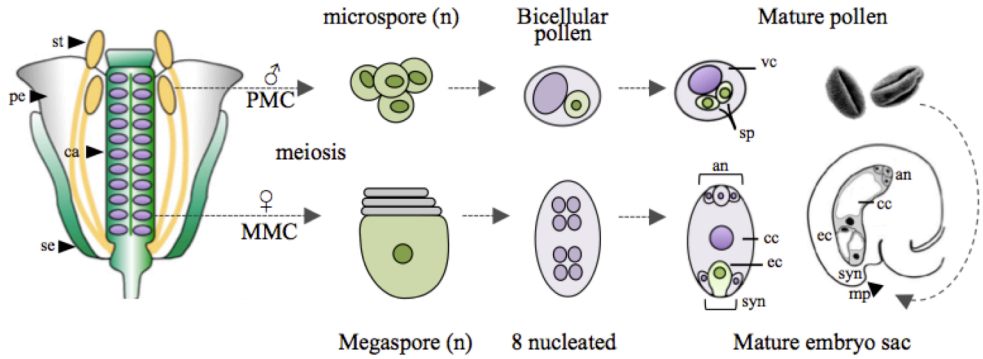


Figure 1.9. Arabidopsis Gametogenesis. Inside the flower, pollen mother cells (PMC) and megaspore mother cells (MMC) are generated in a position-dependent manner from somatic cells in the male and female reproductive tissues. Around flower stage 9, PMC separates from each other, meiosis takes place and generates four microspores (tetrads, 1n). At stage 11, first mitotic division occurs (symmetrical) to give rise to the vegetative cell (vc) and the generative cell. A second mitotic division takes place at stage 12 to create two sperm cells (sp, 1n) within the vegetative cell, which leads to the mature pollen. At stage 13 the pollen is dehiscent. Conversely, at flower stage 10 MMC enlarge and meiosis takes place at stage 11, generating 4 megaspores (1n) with only one been functional (1 nucleated embryo sac). At floral stage 12, three rounds of nuclear divisions generate a syncytial female gametophyte with eight nuclei. Then, cytokinesis takes place at stage 13 to establish the mature female gametophyte, which consists of 8-nucleated/7-celled embryo sac: the egg cell (ec, 1n), the central cell (cc, 2n) and accessory cells, the antipodals (an) and synergids (syn). At flower stage 14 fertilization takes place: pollen tube growths and deliver the 2 vegetative cells trough the micropylar pole (mp) of the ovule, and embryo development begins. (Smyth *et al*., 1990; Bowman, 1994; Schneitz *et al*., 1995; Alvarez-Buylla *et al*., 2010; Kawashima and Berger, 2014).

During plant sexual reproduction a chromatin 'resetting' is carried out to restore pluripotency in the zygote (Kohler and Springer, 2017). Plant germlines possess some mechanisms to bypass possible modifications made by viruses or translocation of transposable elements (TEs). Germlines cells are marked by activation of noncoding RNA-related mechanisms that contribute to genome stability though ARGONAUTE family genes. In Arabidopsis, in particular AGO9 preferentially associates with 24-nt siRNA targeting TEs, but also with some 20– 2-nt mi-RNAs species at the embryo sac (Olmedo-Monfil *et al.*, 2010). Small RNA pathways, including miRNA and tasiRNA pathways, also operate in the male gametophyte (Berger and Twell, 2011; He *et al.*, 2011; Kawashima and Berger, 2014). Vegetative nucleus reprograming leads to TE silencing in the neighbor sperm cells, where an accumulation of 21 nucleotide siRNAs produced after TE activation in the vegetative cell is detected. These siRNA keep DDM1 activity high in the gametes, and prevent transposition of TEs and promoting transgenerational TE silencing (Slotkin *et al.*, 2009). Nonetheless, different studies had questioned the possibility of siRNA travelling from the vegetative to the sperm cells, in additionally to the controversy of the presence of cytoplasmic connection between sperm a vegetative cells, pointing that maybe these siRNAs are inherited from microspores (Kawashima and Berger, 2014).

Sperm and vegetative cells also display different DNA methylation levels during development. CHH methylation decreased from microspores to sperm cells, although CG methylation remains constant. Additionally, in the vegetative nucleus CG methylation is also decreased and H3K9me3 is barely detectable, which likely produce that hypomethylated TEs to become mobile (Slotkin *et al.*, 2009; Calarco *et al.*, 2012; Kawashima and Berger, 2014). Epigenetic reprograming including alteration of DNA methylation by repression of DNA METHYLTRANSFERASE-1 (MET1) at the end of female gametophytic development might cause decondensation of chromatin that affects both central cells and egg cells (Jullien and Berger, 2010). Additionally, low chromatin compaction together with dispersed heterochromatin foci has been observed in central cell and egg cell when compared to accessory cells (Baroux *et al.*, 2011), though the establishment of epigenetic dimorphism between the two female gametes is observed by differencing H3K9me2 patterns between them (Pillot *et al.*, 2010; Baroux *et al.*, 2011).

1.4.4 Embryo

Once female and male gametophytes are mature and flowers fully developed, fertilization process begins. Inside the ovule synergid cells secrete pollen tube attractants and serve as the first point of contact with pollen tubes that will ultimately deliver two sperm cells. In the process of fertilization, one of two synergid cells is penetrated by the pollen tube resulting in synergid cell death. One sperm cell (1p) fuses with the egg cell (1m) and the other fuses with the polar nuclei of the central cell (2m). The fertilized 2n egg cell (zygote: 1m:1p) and 3n central cell (2m:1p) develop into the embryo and endosperm, respectively. These two products of fertilization develop coordinately with the maternal ovule integuments, which give rise to the seed (**Fig. 1.10**) (Lawit *et al.*, 2013; Baroux and Grossniklaus, 2015).

After fertilization, the zygote rapidly elongates and divides asymmetrically to produce an apical and a basal cell, which give rise to the embryo proper and the suspensor, respectively. Divisions of the basal cell produce a file of 8–12 cells forming the suspensor and the apical cell divides symmetrically three times to produce an octant-stage embryo. From there on, distinctly oriented cell divisions shape the radial pattern: at the globular stage protoderm, ground tissue, and provasculature elements are formed from the periphery to the center of the embryo. Morphogenesis, with oriented divisions and anisotropic cell elongation, forms first a heart- and then a torpedo-shaped embryo that comprises the basic tissue patterns and meristematic stem cells that generate the future root and shoot organs (Baroux and Grossniklaus, 2015). Once the mature seed perceive ideal external conditions to germinate,

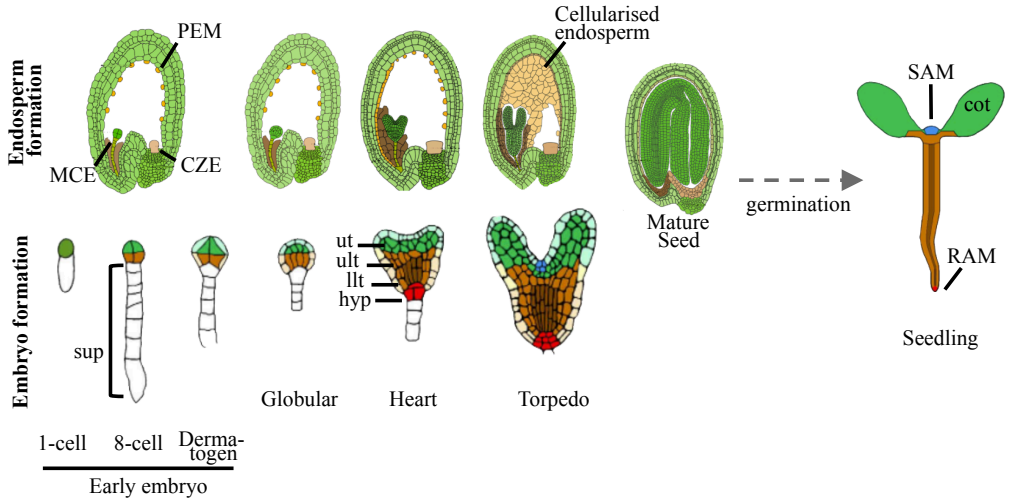


Figure 1.10. Stages of Arabidopsis embryogenesis. After fertilization, the first asymmetric division of the zygote (lower panel), produces an apical cell (green), from which the embryo is mainly derived, and a basal cell, which gives rise to the suspensor and part of the root. After a series of longitudinal, transversal and tangential divisions, above and below tiers are distinguish at heart stage. Colours identify corresponding regions of embryo and seedling. sup: suspensor; ut: upper tier; ult: upper lower tier; llt: lower-lower tier; hyp: hypophyseal cells. Contrariwise, the development of the triploid endosperm (upper panel) is initiated by nuclear divisions (syncytium) at the embryo sac; cellularization commences at late globular stage, and is completed by the upturned-U stage (between torpedo stage and mature seed, not shown). As the cotyledons develop, endosperm is absorbed, leaving only a single layer of it in the matured seed. PEM: peripheral endosperm; MCE: micropylar endosperm; CZE: chalazal endosperm; cot: cotyledons; SAM: shoot apical meristem. Adapted from Capron et al., 2009 and Belmonte et al., 2013.

the radicle and cotyledons will emerged and a new plant will form.

During embryonic development cell cycle genes are upregulated whereas they get repressed in the transition from heart to torpedo stage, where embryos show a high level of functional differentiation in roots and cotyledons (Spencer *et al.*, 2007). This is also visualized by a change in the ratio of canonical histone H3.1 to the variant H3.3, which is high at the heart stage and much lower at the torpedo stage (Otero *et al.*, 2016). This reflects a reduction on the proliferation potential of the cells and the acquisition of cell fate, clearly visualized by WOX5 expression in the QC cells (Otero *et al.*, 2016). Epigenetic reprograming of DNA methylation at specific loci occurs, and reprogramming of histone modifications such as H3K27me2 and H3K27me3 might influence the transition between the gametophytic life and the sporophytic life (Kawashima and Berger, 2014).

In addition to the embryo formation, the endosperm, devoted to nourish the developing embryo, is also controlled by epigenetic effects that influence seed's development (Berger, 1999); Bemer and Grossniklaus, 2012). Endosperm genome is largely hypomethylated compared to that of the embryo (Gehring *et al.*, 2009; Kohler and Kradolfer, 2011). Moreover, interplay between maternal inherited histone H3K27me3 modifications and maternally derived small RNAs is thought to reinforce the epigenetic and transcriptional

dimorphism between the parental genomes controlling endosperm development (Kohler and Kradolfer, 2011; Bemer and Grossniklaus, 2012).

Aims

AIMS

Depending of the organism, reading and interpretation of the histone code can lead to opposite roles in terms of heterochromatin formation and transcriptional activation. In plants, the presence of the trimethylated form of the H4K20 (H4K20me3) has been associated to euchromatic regions, in contrast to what is described in animals. However, the identification of the corresponding methyltransferase has been elusive in *Arabidopsis thaliana*. The general goal of this project is the functional study of Arabidopsis methyltransferases. To accomplish this goal we propose the following specific objectives:

1- Identification of Arabidopsis SET-domain proteins with homology to animal H4K20 methyltransferases

2- Examination of enzymatic activity of candidate SET-domain proteins SDG7 and SDG24

3- Functional analysis of SDG7 and SDG24 in cell dynamics and plant development.

Materials and Methods

3.1. MATERIALS

3.1.1 Plant ecotypes and growth conditions

All T-DNA mutant lines and transgenic lines generated in this study were *Arabidopsis thaliana* Columbia ecotype (Col-0). Plants were grown in an incubator under 21°C, 100 µmol/(m^2·s) light intensity and standard long-day conditions (16 h light /8 h dark) in either or 1% to 0.8% agar MSS plates (1% sucrose, 0.5 x Murashige and Skoog salt, 2.24 g/l Duchefa, and 0,5 g/l MES, pH 5.7). After 7 to 10 days, seedlings were transplanted into soil if needed, and cultivated under the same growing conditions described above in growing chambers.

Arabidopsis MM2d cellular culture (Menges and Murray, 2002) was employed for cell-cycle studies. Cells were kept in dark at 26 ºC with constant agitation at 120 rpm. For maintenance, cells were subcultured every 7 days by diluting 1/20 into new media.

3.2.2 Bacterial strains and growth conditions

Cloning was performed using *Escherichia coli*, either DH5α or Z competent cells (Invitrogen). Gateway vectors were amplified using *E. coli* DB3.1 strain. Recombinant proteins were expressed in *E. coli* DE3 strains BL21 (Stratagene), Rosetta or Rosetta pLysS (Novagen). Plant transformation was carried out with *Agrobacterium tumefaciens*, strain C58C1. Bacterial growths were carried out in liquid or solid (1.5% agar) LB broth (Miller), supplemented according strain specifications.

3.1.3 Bacterial lines expressing recombinant proteins

Table 3.1. Protein bacterial expression lines

Line	Vector	Antibiotic Resistance	Insert
GST-SDG7	pDEST15	Ampicillin	SDG7 CDS
GST-SDG24.1	pDEST15	Ampicillin	SDG24.1 CDS
His6X-SDG7	pDEST17	Ampicillin	SDG7 CDS
His6X-SDG24.1	pDEST17	Ampicillin	SDG24.1 CDS
His6X-E2FC (Desvoyes, unpublished)	pDEST17	Ampicillin	E2FC CDS
GST-ATXR6$_{PHD-SET}$ (Jacob *et al.*, 2009)	pGEX-6P	Ampicillin	ATXR6 CDS (PHD-SET domain)
GST	pEX	Ampicillin	None

3.1.4 Plant lines generated through transformation

Table 3.2. Transgenic plant lines

Line	Vector	Antibiotic resistance	Insert
pSDG7:SDG7-GUS/ CFP/ mRFP	pGW433 (GUS) pGW443 (ECFP) pGW453 (mRFP)	Spectinomycin (Bacteria) Kanamycin (Plant)	2075 kb *SDG7* promoter + *SDG7* genomic sequence without Stop Codon
pSDG24:SDG24.1-GUS/ mRFP	pGW433 (GUS) pGW453 (mRFP)	Spectinomycin (Bacteria) Kanamycin (Plant)	2016 kb *SDG24* promoter + *SDG24.1* genomic sequence without Stop Codon
pSDG24:SDG24.2: GUS	pGW433 (GUS)	Spectinomycin (Bacteria) Kanamycin (Plant)	2016 kb *SDG24* promoter + *SDG24.2* genomic sequence without Stop Codon

3.1.5 Mutant lines

sdg7-2 (Salk_131218)
sdg7-4 (Salk_143603)
sdg7-5 (WiscDLox430F09)
sdg24-1 (WiscDsLOX489-492G18)
sdg24-2 (Salk_201808)
sdg24-3 (Salk_205054)
sdg24-4 (SK22803)
atxr5/6 (Salk_130607 x Sail_240_H01;(Jacob *et al.*, 2009))

3.1.6 Other lines

pORC1b:ORC1b-GFP/ mRFP (Vergara, 2017).
pHTR5:HTR5-GFP (Otero *et al.*, 2016).
pCYCB1;1:CYCB1;1-GFP (Ubeda-Tomas *et al.*, 2009).
pCYCB1;1-GFP (Colon-Carmona *et al.*, 1999).
pWOX5-GFP (Sarkar *et al.*, 2007).

3.1.7 Lines generated by cross pollination

Table 3.3. Crossed lines

Name	Ovule (♀)	Pollen (♂)
sdg7/sdg24	*sdg24-2*	*sdg7-5*
atxr5/6 sdg7	*atxr5/6*	*sdg7-5*
atxr5/6 sdg24	*atxr5/6*	*sdg24-2*
sdg7-5 pWOX5-GFP	*sdg7-5*	pWOX5-GFP
sdg24-2 pWOX5-GFP	*sdg24-2*	pWOX5-GFP
sdg7-5 pCYCB1;1-GFP	*sdg7-5*	pCYCB1;1-GFP
sdg24-2 pCYCB1;1-GFP	*sdg24-2*	pCYCB1;1-GFP
SDG24.1-mRFP ORC1b-GFP	pSDG24:SDG24.1-mRFP	pORC1b:ORC1b-GFP
SDG24.1-mRFP CYCB1;1-GFP	pSDG24:SDG24.1-mRFP	N-CYCB1;1:CYCB1;1-GFP

3.1.8 Oligonucleotides

Table 3.4. Cloning Primers

Amplicon name	Oligo name	Sequence 5′ to 3′	Purpose
SDG7	ASHH3 attB1-0 F	5′- GGGGACAAGTTTGTACAAAAAAGCAGGC-TCTGAGTTCTCATCCGCAGA	Gateway cloning of *pSDG7:SDG7*
SDG7	ASHH3 attB2-aR	5′-GGGGACCACTTTGTACAAGAAAGCTGGGTAGA CAATCTCCCAGTCTTCT	Gateway cloning *pSDG7:SDG7*
SDG24.1 *SDG24.2*	ASHH4 4.5 F -2016	5′- ATAAGTTGTGTCTCCAATAGGCT	Cloning *pSDG24:SDG24.1* or *SDG24.2*
SDG24.1	ASHH4 4.5 R-STOP	5′- ACCAGCAATGCGTCTCCC	Cloning *pSDG24:SDG24.1*
SDG24.2	ASHH4 4.18 R-STOP	5′- AAACTCATCATATTGTGTAATGA	Cloning *pSDG24:SDG24.2*
SDG24.1/ SDG24.2	ASHH4 attB1-0 4.5F	5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTA-TAAGTTGTGTCTCCAAT	Gateway cloning *pSDG24:SDG24.1* or *SDG24.2*
SDG24.1	ASHH4 attB2-a 4.5 R	5′- GGGGACCACTTTGTACAAGAAAGCTGGGTA-ACCAGCAATGCGTCTCCC	Gateway cloning *pSDG24:SDG24.1*
SDG24.2	ASHH4 attB2-a 4.18 R	5′- GGGGACCACTTTGTACAAGAAAGCTGGG-TAAAACTCATCATATTGTGT	Gateway cloning *pSDG24:SDG24.2*
SDG7	ASHH3 clone L	5′- ATGCCAGCCAGCAAAAAGA	Cloning of *SDG7* CDS
SDG7	ASHH3 clone R	5′- TTAGACAATCTCCCAGTCTTCTCT	Cloning of *SDG7* CDS
SDG7	attB1-ASHH3 F	5′- GGGGACAAGTTTGTACAAAAAAGCAGGC-TTCATGCCAGCCAGCAAAAAGA	Gateway cloning of *SDG7* CDS
SDG7	attB2-ASHH3 R	5′- GGGGACCACTTTGTACAAGAAAGCTGGG-TTTAGACAATCTCCCAGTCT	Gateway cloning of *SDG7* CDS
SDG24	ASHH4 clone L	5′- ATGTCTTCCTCGAAGAAGGG	Cloning of *SDG24* CDS variants
SDG24	ASHH4 clone R	5′- TCAAGCAGTTACAACCTTCC	Cloning of *SDG24* CDS variants
SDG24.1	attB2-ASHH4-F2	5′- GGGGACAAGTTTGTACAAAAAAGCAGGC-TTCATGTCTTCCTCGAAGAAGG	Gateway cloning of *SDG7* CDS
SDG24.1	attB2-ASHH4 R	5′- GGGGACCACTTTGTACAAGAAAGCTGGG-TTCAAGCAGTTACAACCTTC	Gateway cloning of *SDG24.1* or *SDG24.2* CDS

Table 3.5. qPCR Primers

Amplicon name	Oligo name	Sequence 5′ to 3′	Purpose
SDG7	ASHH3 3-4 L SP	5′- TGTTCTTGCTCCTCCTCGTC	qPCR for transcript levels
SDG7	ASHH3 3-4 R SP	5′- AGCTTGAGAATAGCATCCCACA	qPCR for transcript levels
SDG7	ASHH3 5-6 L JN	5′- ATGCGGATCAGGGATTGTGG	qPCR for transcript levels
SDG7	ASHH3 5-6 R JN	5′- TCCAAAGCCTCTCTTCACAAGT	qPCR for transcript levels
SDG7	ASHH3 10-11 L SP	5′- CAAAGGGATACATGTTCGCGT	qPCR for transcript levels
SDG7	ASHH3 10-11 R SP	5′- ACGAGGCCAAAACACCTATCA	qPCR for transcript levels
SDG24	ASHH4 1-2 L sp	5′- TTCCTCGAAGAAGGGTTCCG	qPCR for transcript levels
SDG24	ASHH4 1-2 R sp	5′- TCTCCAATTCCCCAATCTGCT	qPCR for transcript levels
SDG24	ASHH4 3-4 L sp	5′- GCGATTGCAATTGTGGGATTCT	qPCR for transcript levels
SDG24	ASHH4 3-4 R sp	5′- GAACGGCTTGTTGGTGCATT	qPCR for transcript levels
SDG24	ASHH4 06-07 L jn	5′- TCAATCACAGCTGCAGTCCT	qPCR for transcript levels
SDG24	ASHH4 06-07 R jn	5′- AGTCAACTGCTCGCCTTTGT	qPCR for transcript levels
ACTIN	ACT2-F	5′- ACGGTAACATTGTGCTCAGTG GTGG	qPCR for transcript levels
ACTIN	ACT2-R	5′- TTGGAGATCCACATCTGCTGGAATG	qPCR for transcript levels
GAPDH	GAPC2 2F	5′- TTGCTCCTCTTGCCAAGGTTA	qPCR for transcript levels
GAPDH	GAPC2 2R	5′- GGACAGTGGTCATGAGTCCC	qPCR for transcript levels
UBIQUITIN	UBQ10 F	5′- GGCCTTGTATAATCCCTGATGAATAAG	qPCR for transcript levels
UBIQUITIN	UBQ10 R	5′- AAAGAGATAACAGGAACGGAAACATAGT	qPCR for transcript levels
BRCA1	BRCA1-F	5′- GTGACGAGCATAAACCCTGATT	qPCR for transcript levels
BRCA1	BRCA1-R	5′- ATGAAATGAGAATACGAAGAAA	qPCR for transcript levels
PARP1	PARP1-F	5′- TGGTTGAAGCATTAGGTGAGA	qPCR for transcript levels
PARP1	PARP1-R	5′- TTCGCAGACCTTGAGATAAAAT	qPCR for transcript levels
RAD51	RAD51-F	5′- TAATAATAAAAATGAGGAGGAT	qPCR for transcript levels
RAD51	RAD51-R	5′- AATGGCGAAACTGAAAACAAAT	qPCR for transcript levels

Table 3.6. Mutant genotyping Primers

Amplicon name	Oligo name	Sequence 5′ to 3′	Purpose
sdg7 T-DNA	Salk131218_LP	5′- TCTCAGATCCGCCATTGTTAC	*sdg7-2* genotyping
sdg7 T-DNA	Salk131218_RP	5′- TCCAAAGCCTCTCTTCACAAG	*sdg7-2* genotyping
sdg7 T-DNA	Salk143603_LP	5′- ACGCAACAAACAATGTCACTG	*sdg7-4* genotyping
sdg7 T-DNA	Salk143603_RP	5′- AAAAAGGGAGAGCATTTGACC	*sdg7-4* genotyping
sdg7 T-DNA	Wisc430F09_LP	5′- AAAGCTTCATCCGAGGCTATC	*sdg7-5* genotyping
sdg7 T-DNA	Wisc430F09_RP	5′- TGATCAAGATTTGTTTTGGCC	*sdg7-5* genotyping
sdg24 T-DNA	Wisc492G18_LP	5′- AGGTAGAGGAAGTGGTGGTGG	*sdg24-1* genotyping
sdg24 T-DNA	Wisc492G18_RP	5′- CCAATTTCTTCCACACCAAAG	*sdg24-1* genotyping
sdg24 T-DNA	Salk201808_LP	5′- TTTGGCTCGATTATCGACAAG	*sdg24-2* genotyping
sdg24 T-DNA	Salk201808_RP	5′- GCACAGCATTTCTGTTATCACG	*sdg24-2* genotyping
sdg24 T-DNA	Salk205054_LP	5′- CCGTTCCAACAACGACATATC	*sdg24-3* genotyping
sdg24 T-DNA	Salk205054_RP	5′- CTTTGGTGTGGAAGAAATTGG	*sdg24-3* genotyping
sdg24 T-DNA	SK22803_LP	5′- TGTGTTAGGACTGCAGCTGTG	*sdg24-4* genotyping
sdg24 T-DNA	SK22803_RP	5′- TCTTTCCCAATTAATCCGGTC	*sdg24-4* genotyping
atxr5 T-DNA	ATXR5 LP	5′- CCATTGGAACTTGGCTTTGTGTC	*atxr5* genotyping
atxr5 T-DNA	ATXR5 RP	5′- AATAGGACCATCTGCTTCAACTGTG	*atxr5* genotyping
atxr6 T-DNA	ATXR6 LP	5′- AGCTTTGCTGGTTGTTTACCGGA	*atxr6* genotyping
atxr6 T-DNA	ATXR6 RP	5′- CCATGTTGAGTAAATGTCGAAGAC	*atxr6* genotyping

Table 3.7. rDNA Primers

Amplicon name	Oligo name	Sequence 5′ to 3′	Purpose
45S 3′ETS rDNA	3ETS var region F	5′- GACAGACTTGTCCAAAACGCCCAC	Semi qPCR for rDNA
45S 3′ETS rDNA	3ETS var region R	5′- CCTGGTCGAGGAATCCTGGACGATT	Semi qPCR for rDNA
18S rDNA	18S 488bp F	5′- GCTCGAAGACGATCAGATAC	Semi qPCR for rDNA
18S rDNA	18S 488bp R	5′- AGACCTGTTATTGCCTCAAA	Semi qPCR for rDNA
5S rDNA	OL10	5′- CCTCGTGTTGCATCCCTC	Semi qPCR for rDNA
5S rDNA	OL9	5′- CTTCCCGGGAGGTCACCC	Semi qPCR for rDNA
ACT2	Actin for 3ETS F	5′- GAGAGATTCAGATGCCCAGAAGTC	Semi qPCR for rDNA
ACT2	Actin for 3ETS R	5′- TGGATTCCAGCAGCTTCCA	Semi qPCR for rDNA
CDT1a	CDT1a-F	5′- ACTGAACAAGCAGCATAAAC	Semi qPCR for rDNA
CDT1a	CDT1a-R	5′- TCTTCTTCTTCCGTCGTCTT	Semi qPCR for rDNA

3.1.9 Antibodies list

Table 3.8. Primary antibodies

Antibody	Reference	Working dilution
α-GST	# sc-138 (Santa Cruz Biotechnology)	WB 1/3000
α-poly HIS	# H1029 (SIGMA)	WB: 1/5000
α-RFP	# R10367 (Thermo-Fischer)	IHC: 1/500-1/250
α-DsRed (E-8)	# sc-390909 (Santa Cruz Biotechnology)	WB: 1/500 IHC: 1/250 1/500
α-mCherry	# DSHB-mCherry-3A11 (Developmental Studies Hybridoma Bank)	IHC: 1:2000
α-H3K4me2	# 39141 (Active motif)	WB: 1/12000
α-H3K4me3	# ab8580 (Abcam)	WB: 1/6000
α-H3K9me2	# ab1220 (Abcam)	WB: 1/5000
α-H3K9me3	# 39161 (Active motif)	WB: 1/20000
α-H3K27me1	# 07-448 (Merck- Millipore)	WB: 1/5000
α-H3K27me2	# 39245 (Active motif)	WB: 1/5000
α-H3K27me3	# 39535 (Active motif)	WB: 1/2000
α-H3K36me3	# ab9050-100 (Abcam)	WB: 1/10000
α-H4K20me1	# ab9051 (Abcam)	WB: 1/1000
α-H4K20me1	#39728 (Active Motif)	WB: 1/200000
α-H4K20me2	# 39173 (Active Motif)	WB: 1/6000
α-H4K20me3	# 07-463 (Merck- Millipore)	WB: 1/500 -1/1000
α-H4K20me3	# ab9053 (Abcam)	WB: 1/500
α-H4K20me3	#AR0136-200LP Bio	WB: 1/1000
α-Orc1b	Particular Serum (Vergara, 2017)	WB: 1:5000

3.2. METHODS

3.2.1 Molecular biology Techniques

3.2.1.1 DNA extractions

Genomic DNA extraction was done by the CTAB method (Doyle and Doyle, 1990),with some modifications. Roots, leaves, flowers and seeds were frozen in liquid nitrogen together with glass beads and ground in Silamat S5 (Ivoclar Vivadent) for 10 seconds. In particular, seeds were homogenated with the help of a mortar and pestle in liquid nitrogen. Next, 500 µL of extraction buffer (100 mM Tris-HCl, pH 8.0, 1.4 M NaCl, 20 mM EDTA, 3% CTAB, 140 mM β-mercaptoethanol) was preheated at 65 ºC and added to the tissue. Samples were once again homogenated in Silamat for 5 seconds and heated at 65 ºC for 15- 20 minutes. After incubation, 500 µL of chloroform: isoamyl alcohol (24:1) was added; then, the mixture was shaken and centrifuged at maximum speed (20000 x*g*) for 10 minutes. The supernatant was collected and DNA precipitated by adding 810 µL of 100% ethanol and 1/10 3 M of sodium acetate (pH 5.2), incubation proceeded at 4 ºC for 20-30 minutes. The DNA pellet was collected by centrifuging at maximum speed for 15 minutes. The pellet was washed with 70% ethanol, air-dried for at least 30 minutes and resuspend in 40 µL of water. DNA was quantified by using NanoDrop (Thermo-Fischer Scientific).

DNA extractions were also performed with nuclei obtained by fluorescence-activated cell sorting (FACS). Nuclear membranes were dissolved in Nuclei Lysis Buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS) (Villar and Kohler, 2010). Then, DNA was extracted with 1 volume of phenol: chloroform: isoamyl alcohol (25:24:1) and centrifuged with maximum speed for 15 minutes at 10 ºC. The pellets were washed with 70% ethanol, centrifuged at maximum speed, air- dried for at least 30 minutes and resuspended in 20 µL of water.

For plasmid DNA extraction, bacterial pellets were treated with Wizard Plus SV minipreps purification kit (Promega) following manufacture's instructions.

3.2.1.2 Total RNA extractions

Seedlings, roots, leaves and flowers grown for the appropriate time were frozen together with glass beads in liquid nitrogen. Trizol (250 µL, Thermo-Fischer) was added to the tissue and ground for 8 seconds in Silamat S5. Another 250 µL of Trizol were added and mixed thoroughly. A 100 µL of chloroform were used to extract the total RNA. After spinning the samples for 10 minutes at 4 ºC with maximum speed, the upper phase was transferred to

another tube. The RNA was precipitated using 1 volume of isopropanol for 15 min at -20 ºC and then centrifuged for 10 minutes at 4 ºC. The pellet was washed with 75% ethanol, air dried for 20-30 minutes and then resuspended in 44 µL of nuclease free water. After that, RNA was treated with 0.5 µL of DNase I (Roche, 10 U/µL) for 20 minutes at 37 ºC and extracted again using phenol: chloroform: isoamyl alcohol (25:24:1). RNA was then precipitated with 1/10 3 M sodium acetate (pH 5.2) and 2.5 volumes of 100% ethanol. The pellets were then washed with 75% ethanol, air dried and resuspended in 40 µL of nuclease free water. The RNA was quantified using NanoDrop, and its quality was assessed by fractionating 100-500 ng of RNA in a 1.2% agarose gel stained with RedSafe (1 µL/100 mL). Bands were subjected to UV light detection.

RNA extraction from seeds was done following the method described by Oñate-Sánchez and Vicente-Carbajosa (2008). Seeds were frozen in liquid nitrogen and ground by mortar and pestle. Samples were transferred to an Eppendorf tube and added 550 µL of cooled extraction buffer (0.2 M Tris-HCl, pH 8.0, 25 mM EDTA, 0.4 M LiCl, 1% SDS) and 550 µL of chloroform. The mixture was vortexed vigorously for 10 seconds and centrifuged for 3 minutes; then, the supernatant was transferred to a new tube. Next, 500 µL of water-saturated acidic phenol was added, vortexed vigorously, mixed with 200 µL chloroform and centrifuged for 3 minutes. The supernatant was collected and 1/3 volume of 8 M LiCl was added. Precipitation proceeded by incubating the sample at -20 ºC for 1 hour and centrifuging with maximum speed for 30 minutes at 4 ºC. The pellet was dissolved in 26 µL of nuclease free water and treated with 1 µL of DNase I (Roche, 10 U/µL) for 30 minutes at 37 ºC. Carbohydrates were precipitate by adding 470 µL of water, 7 µL of 3 M sodium acetate (pH 5.2), 250 µL of 100% ethanol and centrifuging for 10 minutes at 4 ºC. The supernatant was collected and the RNA was precipitated by adding 43 µL of 3 M sodium acetate (pH 5.2) and 750 µL of 100% ethanol. The mixture was incubated at -20 ºC for 1 hour and centrifuged for 20 minutes at 4 ºC. Finally, the pellets were washed with 70% ethanol, air dried and resuspended in 20 µL of nuclease free water.

3.2.1.3 Protein extractions

Extraction of nuclear proteins from MM2d cells and plants was done according to the protocol described in Villar and Köhler (2010), with some modifications. 1-3 g of tissue was frozen in liquid nitrogen and homogenized to a fine powder with the aide of a mortar and pestle. For each 1 g of tissue, 10 mL of extraction buffer 1 (EB1: 10 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 0.4 M sucrose, 0.1 mM PMSF, 1X plant protease inhibitors cocktail (Sigma)) was added into the frozen material and mixed by inversion until powder dissolved. The

solution was filtered through a double layer of fine nylon mesh and passed twice through a dounce homogenizer (loose and tight). The nuclei solution was then centrifuged at 3000 x*g* for 20 minutes at 4 ºC and supernatant was discarded. The pellet was dissolved in 1 mL of extraction buffer 2 (EB2: 10 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 0.25 M sucrose, 1% Triton X-100, 0.1 mM PMSF; 1X plant protease inhibitors cocktail (Sigma)) and centrifuged at 12000 x*g* for 10 minutes at 4 ºC. The supernatant was discarded and nuclei were resuspended in 150-200 µL of nuclei extraction buffer (NEB: 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1X plant protease inhibitors cocktail (Sigma)). The nuclei solution was then sonicated 5 times 10 s ON/ 30 s OFF cycle, on a Bioruptor. SDS was added after sonication to a final concentration of 1% and mixed carefully by pipetting up and down. Dissolved nuclear proteins were quantified with the Pierce BCA protein assay (Thermo-Fischer) according to the manufacturer's instructions. Nuclear protein extractions were evaluated by SDS-PAGE on 10-12% Tris-glycine (mRFP tagged proteins) or 15% Tris-Tricine gels (histones), transferred to a PVDF membranes (0.2 µm pore size, GE) and detected by Western blot using specific antibodies.

To measure the level of methylation marks in mutant lines, protein extractions were carried out using several protocols in order to obtain higher histones ratio in nuclear fractions. Different starting amounts of tissue (1- 5 g) were frozen in liquid nitrogen and homogenized to a fine powder with the aide of a mortar and pestle. In a first attempt, a protocol described earlier for MM2d cells was used with some additional steps (Lin Xu, National Laboratory of Plant Molecular Genetics, Shanghai- China, Personal communication). Nuclei were extracted with EB1 buffer, filtered trough a nylon mesh and homogenated with the help of a dounce homogenizer. Then, the nuclei were centrifuged and the pellet was dissolved in EB2 buffer. Consecutively, histones were extracted from the pellet by adding 200 µL of 0.4 M H_2SO_4 and incubated for 45 minutes with soft rotation. The solution was centrifuged at 10000 x*g* for 5 minutes at 4 ºC, the supernatant collected, and histones were precipitated in two ways: (i) by ethanol or (ii) by TCA precipitation. In the first case, 5 volumes of 100% ethanol were added and the mix was incubated for 48 hours at -20 ºC, centrifuged at 20000 x*g* for 10 minutes at 4 ºC, and the pellets were washed with cooled 80% ethanol. In the second case, 20% TCA (trichloroacetic acid) was added to the protein mix, incubated over night at 4 ºC, centrifuged at 20000 x*g* for 10 minutes at 4 ºC, and the pellets were washed with cold acetone. Final pellets for both cases were air dried and resuspended in 50 µL of water over night. Histones were quantified by the BCA assay and tested by SDS-PAGE and Western blot, as described above.

Other additional protocols were tested including: (i) nuclei extraction described by Deal and Henikoff (2010), and (ii) histone high-salt extraction protocol described in Shechter et al. (2007).

3.2.1.4 Genotyping PCRs

The Biotools DNA Taq Polymerase kit was used according to the manufacturer's instructions. Briefly, in an Eppendorf tube were mixed 2.5 µL of 10X reaction buffer supplemented with $MgCl_2$, 0.5 µL of dNTPs mix (10 mM each), 0.5 µL of primer mix (10 µM each), 0.5 U of Taq polymerase (1 U/µL) and 1 µL of DNA in a final volume reaction of 25 µL. The PCR program was set as follows: one cycle of 3 minutes at 94 ºC; 35 cycles of 94 ºC for 30 seconds, 55 ºC for 30 seconds and 72 ºC for 1 minute; and one final extension of 72 ºC for 5 minutes. The PCR products were resolved in 1% agarose gels containing 1 µL of RedSafe per 100 mL of buffer and visualized with UV light.

3.2.1.5 RT- PCRs

To obtain cDNA, 500-1000 ng of total RNA was used. The reverse transcriptase reaction was performed using oligo-dT as a primer and the SuperScript III kit from Thermo-Fischer following the kit's instructions. For CDS amplification of *SDG7* and *SDG24* specific primers were used. For rDNA analysis, cDNA was amplified by using random primer hexamers. Complementary RNA was removed by treating samples with 0.4 Units of RNase H (2 U/ µL) for 20 minutes at 37 ºC.

3.2.1.6 Quantitative PCR (qPCR)

Depending on the region to amplify, cDNA was diluted between 4-16 times, and 2 µL were mixed with the desired primers and the reagents of the GoTaq qPCR master mix (Promega). Standards, negative controls and samples were run in ABI Prism 7900HT SDS. Primer's T_m was designed to perform at 60 ºC. The efficiency of amplification for each primer pair was determined setting a standard curve with different dilutions of cDNA. The slope of the line (n) was used to calculate the efficiency with the following formula: $10*^{(-1/n)}$. To calculate the relative expression of different genes, the $2^{-\Delta\Delta Ct}$ method was used. First, the Cts of the gene of interest (target) and the Cts of the reference gene (*ACT2*, *GAPDH2* or *UBQ*) were compared ($\Delta Ct = Ct_{target} - Ct_{reference}$). Then, the ΔCt of each sample was normalized to the control sample Col-0 ($\Delta\Delta Ct = \Delta Ct_{target} - \Delta Ct_{control}$).

3.2.1.7 Semi-quantitative PCR (sqPCR)

DNA or cDNA templates (100 ng) were used for PCR amplification of rDNA sequences (45S 3'ETS, 18S and 5S) or controls (*ACT2* and *CDT1a*). 0.5 U of Taq polymerase (1 U/µL, Biotools) was used in each PCR reaction. Several PCR amplification cycles were applied (16, 20, 25, 27 or 30) with the following conditions: 30 seconds at 94 ºC, 30 seconds at 56 ºC, and 1 minute at 72 ºC. sqPCR products were resolved by electrophoresis on 2.5% agarose gels (stained with RedSafe) during 2 hours at 110 V.

3.2.1.8 Sequencing

PCR products or purified plasmid constructions were subjected to electrophoresis in 0.8% agarose gels with RedSafe. Bands were excised with the aid of a razor bladed under UV light, purified with Wizard PCR Clean-Up System (Promega) and quantified by NanoDrop. DNA was mixed with the appropriate primer and sent for sequencing to MACROGEN. Chromatograms and sequences were analyzed with SeqBuilder and SeqMan Pro software (DNASTART suite).

3.2.1.9 Gateway cloning

To determine the endogenous protein expression of *SDG7* and two *SDG24* variants (*SDG24.1* and *SDG24.2*) in Arabidopsis organs, a genomic fragment containing the putative promoter (~2 kb) and the coding region (except for the termination codon) were amplified by PCR using the KOD DNA polymerase (Merck- Millipore) and specific primers containing the attB sites. The PCR products were purified using Wizard SV Gel and PCR Clean-Up system (Promega) and cloned into pDONR221 using BP Clonase II (Thermo-Fischer Scientific). Colonies were analyzed using restriction enzymes after obtaining plasmid DNA, and those that were positive were sent for sequencing. After sequence confirmation, vectors were subjected to LR reaction with LR Clonase II (Thermo-Fischer Scientific) into destination vectors carrying different C- terminal fusion tags, including GUS (β- glucuronidase), mRFP (monomeric red fluorescent protein), and eCFP (enhanced cyan fluorescent protein). The translation of the *SDG7*, *SDG24.1* and *SDG24.2* tagged lines will only result in the production of those specific products to be in frame with the different tags. Destination vectors were kindly given by Tsuyoshi Nakagawa (Nakagawa *et al.*, 2007). Cloning primers, destination vectors and growing conditions are detailed in Materials.

Cloning of the coding sequences of *SDG7* and *SDG24.1* for recombinant protein expression was done with cDNA templates obtained from seedlings and flowers, respectively.

The cDNA was amplified by PCR using Pfx accuprime DNA polymerase (Thermo-Fischer Scientific) and specific primers containing the attB sites. PCR products were purified and cloned into pDONR221 using BP Clonase II. The resulting vectors were confirmed by sequencing, and recombined with LR Clonase II into destination vectors containing N-terminal fusion tags: GST (Glutathione S-Transferase) or His6X (6 histidine residues) from Thermo-Fischer Scientific. Details of primers, destination vectors and growing conditions are presented in Materials.

3.2.1.10 Recombinant protein expression in bacteria

Vectors carrying *His6X-SDG7* or *His6X-SDG24.1* constructions were transformed into BL21 Rosetta pLysS and plated into LB agar supplemented with chloramphenicol (25 µg/mL) and ampicillin (50 µg/mL) overnight. One single colony was taken to start cultures (30 mL of LB supplemented with respective antibiotics and 0.5% of Glucose). Next day, the starter culture was diluted 1:50 in a pre-warmer LB medium (supplemented with respective antibiotics and 0.25% of glucose) to a final volume of 1 L. Incubation proceeded at 37 ºC with agitation until 0.4 OD of cells was reached. Then, cells were incubated in an ice-water bath for 7 minutes and induced with 0.3 mM of IPTG for 2.5 hours at 18 ºC with agitation. Cell pellets were recovered by centrifuging at 5000 x*g* for 15 minutes at 4 ºC. Then, pellets were washed with cooled 1X PBS, centrifuged at 5000 x*g* for 3 minutes and frozen in liquid nitrogen.

Purification started by thawing bacterial pellets for 20 minutes and re-suspending them in 20 mL of Lysis buffer (50 mM NaH_2PO_4, pH 8.0, 800 mM NaCl, 1% triton X100, 5 mM imidazole, 1X bacterial protease inhibitors cocktail (Sigma)) by pipetting up and down. Lysozyme was added to the lysate to a final concentration of 0.5 mg/mL and incubated for 15 minutes with soft rotation at 4 ºC. The pellets were sonicated (4 rounds of 10 seconds at high amplitude) and centrifuged at 20000 x*g* for 20 minutes at 4 ºC. Supernatants were collected and incubated for 1-2 hours with 400 µL of Ni-NTA matrix beads (previously washed with Lysis buffer) at 4 ºC with soft rotation. Beads were washed three times with Wash buffer (50 mM NaH_2PO_4, pH 8.0, 600 mM NaCl, 0.1% triton X-100, 10% glycerol, 50 mM imidazole, 1X bacterial protease inhibitors cocktail (Sigma)). At this point, protein bound to beads could be kept on ice for HMTase assay or to be eluted from the beads for Peptide binding assay.

For elution, beads were incubated with 350 µL of Elution buffer (50 mM NaH_2PO_4, pH 8.0, 500 mM NaCl, 0.1% NP40 Igepal, 10% Glycerol, 500 mM Imidazole, 1X bacterial protease inhibitors cocktail (Sigma)) for 30 minutes at 4 ºC with soft rotation. Beads were

centrifuged and the supernatant was collected and kept on ice. Elution step was repeated three more times under the same conditions. Eluted proteins were subjected to dialysis (individual elution steps were kept separated) with 13 kDa-pore membranes in Binding buffer (50 mM Tris-HCl, pH 9.0, 5 mM $MgCl_2$, 0.1 mM PMSF). Incubation proceeded over night at 4 ºC. Next day, dialyzed proteins were submitted to Western blot analysis and Coomassie staining for quantification.

For *in vitro* HMTase assay other proteins were also induced. GST protein was obtained by inducing appropriate construct in a 50 mL BL21 cell culture with 0.5 mM IPTG for 1 hour at 30 ºC. For ATXR6, 200 mL of induced GST-ATXR6 protein (BL21 Rosetta cells) was obtained as describe in Jacob et al. (2009). Both GST and GST-ATXR6 purification steps were also performed as in Jacob et al. (2009) leaving recombinant proteins attached to the glutathione matrix without elution. On the other hand, 300 mL of His6X-E2FC were induced in BL21 cells with 0.5 mM IPTG for 1.5 hours at 30 ºC. Protein purification followed the same conditions as for His6X-SDG7 protein, leaving recombinant protein attached to Ni-NTA matrix.

3.2.1.11 Recombinant protein expression in bacteria

3.2.1.11.1 Histone peptide binging assay

Dialyzed Recombinant proteins were incubated with MODified Histone peptide array (CelluSpots, Active Motif) following the manufacture's instructions. Briefly, arrays were immersed into blocking solution (TTBS buffer: 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05% Tween-20, supplemented with 5% milk) and incubated with soft agitation (orbital shaker) for 1 hour at room temperature. A quick rinse of 30 seconds was performed and followed by 3 washing steps of 5 minutes at room temperature with TTBS buffer. Recombinant proteins were diluted in Binding buffer (50 mM Tris-HCl, pH 9.0, 5 mM $MgCl_2$) supplemented with 4 mM DTT and 1X bacterial protease inhibitors cocktail (Sigma) to a final volume of 3 mL and added to peptide arrays. Incubation proceeded over night at 4 ºC with soft agitation.

Next day, a quick rinse of 30 seconds and 3 washing steps of 5 minutes at room temperature with TTBS buffer were performed. First antibody (α-polyHis) was diluted 1:3000 in blocking solution with a final volume of 6 mL, and incubated with peptide array with soft agitation for 1 hour at RT. A quick rinse of 30 seconds and 3 washing steps of 5 minutes at RT with TTBS buffer were performed. Second antibody (α-mouse HRP) was diluted 1:10000 in 6 mL of blocking solution and incubated with the peptide array with soft agitation for 1

hour at room temperature. Another quick rinse of 30 seconds and 3 washing steps of 5 minutes at room temperature with TTBS buffer were performed. For detection, the ECL developing solution was added to each slide, incubated over 2 minutes and detected with a CCD camera and films during different time points.

The resulting images were re-sized in order to accommodate them to the digital grid given by the Active Motif Analyze Software. The software analyzed spot intensity of the interactions from the array, assigning the number 1 to the most intense spot and 0 (or close) to the less intense signal/background). The software generated an excel file and a graphical analysis for the histone peptide interactions.

3.2.1.11.2 Histone methylation assay

The *in vitro* HMTase assay was essentially performed as described in Jacob et al. (2009) with some modifications. 10 µL of matrix-bound recombinant proteins (His6X-SDG7, His6X-SDG24.1, His6X-E2FC, GST-ATXR6) containing around 3-10 µg of protein were transferred to a micro-centrifuge tube (200 µL); any remaining buffer of the matrix was eliminated by centrifuging for five seconds at 500 *xg*, and pipetted out carefully. Then, on ice, methylation buffer (50 mM Tris-HCl, pH 8.5, 20 mM KCl, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, 250 mM sucrose) and 10 µg of calf thymus histones (Roche) were added to the matrix-bound proteins. The reaction started with the addition of 150 µM of SAM (*S*-(5′-adenosyl)-*L*-methionine chloride dihydrochloride, Sigma). The final volume reaction was set in 40 µL (without taking into account the matrix volume). Mixes were incubated for 2-3 hours at 30 ºC in soft rotation. To avoid disparities in the final histones quantities in each single reaction, methylation buffer and histones were mixed in a master mix and then distributed individually to each reaction.

Reactions were stopped by adding 40 µL of 2X Tricine SDS-PAGE sample buffer followed by heating to 95 °C for 5 min. Tubes were centrifuged for 5 seconds, and 3 µL of supernatant was subjected to SDS-PAGE on 15% Tris-Tricine gels and transferred to PVDF membrane (0.2 µm pore size, GE). Detection was performed by Western blot using specific histone H3 or H4 antibodies. Other methylation conditions were tested as well, by changing methylation buffer and pH conditions (250 mM Tris-HCl, pH 8.5/9.0/9.5, 25 mM $MgCl_2$, 4 mM DTT) and temperature (25 ºC, 37 ºC) (Yannick Jacob, Yale University Connecticut-USA, personal communication).

3.2.2 Cellular biology Techniques

3.2.2.1 Transgenic lines generation

Agrobacterium tumefaciens (C58C1 strain) was transformed and grown for 2 days at 30 ºC in plates containing two antibiotics, rifampicin [50 µg/ml] and the antibiotic of each construct. Then, a bacterial culture of 5 ml was set to grow overnight. An inoculum of 200 ml was centrifuged at 3000 x*g* for 15 minutes and resuspended in a solution of 5% sucrose and 0.05% Silvet l-77. *A. thaliana* Col-0 wild type plants were transformed using the floral dip method (Clough and Bent, 1998) and transformants were selected with 50 µg/mL of kanamycin and 50 µg/mL of timentin in MS plates with 0.6% agar (the addition of timentin to the first generation of transformants avoids Agrobacterium contamination from the collected seeds). In the next generation (T2: transformant F2 generation), plants with only one insertion were selected, and in the T3 homozygous lines, either using antibiotics or looking at the fluorescence in roots with confocal microscope.

3.2.2.2 Seed sterilization

For plant growth, seeds were surface-sterilized in 20% sodium hypochlorite and 0.1% Tween-20 for 8 min and washed four times in sterile water. Seeds were stratified in the dark for 3 d at 4 °C.

3.2.2.3 Quantification of lateral roots (LR) and lateral root primordia (LRP)

For the quantification of total LRPs and emerged LR, seeds were grown for 6 and 10 days, respectively. LR density was scored as the lateral root number per centimeter of primary root and was calculated by dividing the number of LR by the primary root length for each seedling (25-30 seedling were evaluated) (Fernandez-Marcos *et al.*, 2017). For visualization of LRP, seedlings were clarified by fixation in acidic-methanol (20% Methanol; 1.46% HCl) 25 minutes at 55 ºC. After fixation, roots were incubated for 15 minutes in a 7% NaOH/ 60% ethanol solution at room temperature. Roots were transferred to a 50% glycerol solution and kept at 4 ºC until observation at the optical microscope.

3.2.2.4 Genetic crosses

Unfertilized flowers were emasculated by the removal of sepals, petals and stamens with the aid of a sharp microscopy tweezers and a macroscope. Pollination took place by

approaching the pollen in the anthers from a recently opened flower of the other parental line with the help of microscopy tweezers.

For the dissection of the inner flower organs, flowers were first emasculated as described above. Dissection of developing embryo sacs from flower stage 12-13, fertilized embryo sacs and developing embryos was made under the macroscope with the help of syringe needles (BD Microlance 3, 30G 1/2″) and a tape-coated base for immobilization. Samples were placed in 50% glycerol after dissection for observation, or in other buffers according assays requirements.

3.2.2.5 Histochemical staining

3.2.2.5.1 Gus staining

Transgenic lines expressing the GUS tag (β-glucuronidase) were placed in GUS staining solution (100 mM Na_2HPO_4 / NaH_2PO_4, pH 7.0, 1 mM $K_3Fe(CN)_6$, 1 mM $K_4Fe(CN)_6$, 0.1% Tween-20; 1 mM X-Glu dissolved in DMSO) and incubated at 37 ºC. **Table 3.9** summarizes the incubation times and other treatments applied depending on the evaluated tissues. In particular, flowers were subjected to 10 min of vacuum with GUS staining solution prior incubation. For embryos, GUS staining protocol described by Robert *et al*. (2015) was followed. Briefly, embryo sacs or dissected embryos from fertilized flowers were collected in 90% acetone and fixed with cold 90% acetone for 10 min applying vacuum; then, 3 washing step were made with washing solution (50 mM Na_2HPO_4 / NaH_2PO_4, pH 7.0), and the GUS staining solution was infiltrated applying vacuum for 20 min. For chlorophyll elimination, seedlings, leaves and flowers were subjected to subsequent washing steps with 70% ethanol and kept in a 50% glycerol solution at 4 ºC until observation at the optical microscope or a macroscope.

Table 3.9. GUS incubation times

Line	Tissue	Incubation time	Other treatments
SDG7-GUS	Roots and Seedlings	30 min	-
SDG7-GUS	Flowers	1-2 h	10 min vacuum
SDG7-GUS	Embryos	2 h	90% acetone fixation +10 min vacuum.
SDG24.1-GUS	Roots and Seedlings	24 h	
SDG24.1-GUS	Flowers	48-72 h	10 min vacuum
SDG24.1-GUS	Ovules inside Flowers	48-72 h	10 min vacuum
SDG24.2-GUS	Flowers	48-72 h	10 min vacuum

3.2.2.5.2 Root staining for optical and confocal microscopy

To count cortical cell layers and differentiate columella cells from the root, starch grains were stained with a 1% lugol solution by submerging them 1 minute in lugol and washed them with PBS buffer. For root cell wall staining, roots were incubated in propidium iodide dissolved in water (1/50 dilution from 1 mg/L stock). In both cases, observation proceeded immediately at optical or confocal microscopes.

For SDG24.1-mRFP nuclei staining, roots were briefly fixed in 4% paraformaldehyde/ 1X MTSB (Microtubule-stabilizing buffer: 50 mM Pipes, pH 7.0, 5 mM $MgSO_4$, 5 mM EGTA) for 5 minutes and washed twice with 1X MTSB. Then, roots were incubated with 0.25 µg/mL DAPI (4′, 6-diamidino-2-phenylindole) diluted in 1X PBS for 10 minutes, and then washed twice with 1X-PBS buffer before confocal microscope observation. For SDG7-mRFP nuclei observation, nuclei were stained with 0.25 µg/mL DAPI for 5- 10 minutes, and then washed twice with 1X MTSB.

3.2.2.5.3 Ovule staining

Dissected ovules from unfertilized flowers were stained as described in Musielak *et al*., (2015) with SR solution (PBS buffer, pH 8.0, 1% DMSO, 0.05% Triton-X100, 5% glycerol, 4% paraformaldehyde, 0.1% SR2200 (Renaissance Chemicals)). For better penetration, soft vacuum was applied for 5 minutes at RT. Afterwards; the staining solution was replaced with water and again incubated under soft vacuum for 5 minutes. Finally, water was replaced with 10% glycerol solution and ovules were observed under the confocal microscope within 30 minutes after the ovules releasing.

3.2.2.5.4 EdU staining

For detection of S-phase cells in roots, 5 dps seedlings were labeled by adding 50 µM EdU (ethynyl-2′-deoxyuridine) to liquid MSS growth medium during 15 minutes. Then, roots were washed once with 1X MTSB buffer and fixed in 4% paraformaldehyde/ 1X MTSB buffer applying soft vacuum for 20 minutes (protected from light). The fixation solution was removed by washing 4 times with 1X MTSB for 10 minutes, 2 times with 1X PBS for 10 minutes, 2 times with 1X PBS for 5 minutes, and once with water for 5 minutes. Seedlings were placed in charged slides (Thermo scientific Menzel-glaser superfrost plus) and dried overnight. Next day, cotyledons were removed with a razor blade and roots were delimited in the slide with a hydrophobic barrier (PAP pen). Then, root tips were re-hydrated for 10 minutes with 1X MTSB. The cell wall was permeabilized by incubating whit a driselase

solution (20 mg/mL in 1X MTSB) for 40 minutes at 37 ºC in a humid chamber. The driselase solution was removed by washing 4 times with 1X PBS. Afterwards, cell membrane permeabilization took place by incubating root tips with 1X MTSB, 10% DMSO, 3% NP40 (Nonidet P-40 Igepal) solution for 1 hour at room temperature; and then, 3 washing steps with 1X PBS were performed.

Next, EdU was detected by using the Click-iT EdU Alexa fluor 488 imaging kit (Thermo Fisher Scientific) according to the manufacturer's recommendations for 30 minutes. Then, roots were incubated with DAPI (10 µg/mL) for 20 minutes, and then washed once with 1X PBS and once with water. Slides were mounted in Mowiol 4-88 (Sigma) and kept at 4 ºC until confocal microscopy observation.

For EdU labeling compatible with mRFP "*in vivo*" detection, 5 dps seedlings expressing SDG24.1-mRFP were labeled by adding 50 µM EdU to liquid MSS growth medium and incubating for 15 minutes. Then, roots were fixed with 4% paraformaldehyde/ 1X MTSB buffer for 5 minutes, washed twice with 1X MTSB, and permeabilized with 3% NP40 for 10 minutes. Roots were then washed twice with 1X MTSB and placed in a regular microscopy glass slide, and a hydrophobic barrier area was drawn. Roots were immobilized temporally with a small piece of tape by the cotyledons. Immediately, EdU was detected by using the Click-iT EdU Alexa fluor 647 imaging kit (Thermo Fisher Scientific) according to the manufacturer's recommendations for 30 minutes. Roots were DAPI stained (10 µg/ mL) for 10 minutes and washed once with 1X PBS and once with water. The tape was removed and slides were mounted with 10% glycerol solution and observed at the confocal microscope directly.

3.2.2.5.5 EdU labeling combined with immunostaining of mRFP-tagged proteins

Seedlings expressing mRFP-tagged proteins were immunodetected by mRFP specific antibodies following the protocol described in Vergara (2017) . Firstly, seedlings were EdU-labeled and permeabilized as described above for EdU detection in mutant lines. After that point, roots were treated with blocking solution (3% BSA; 10% horse serum; 1X PBS) and incubated for 1 hour at 37 ºC in a humid chamber. Then, primary antibodies (α-RFP, α-DsRed or α-mCherry) were diluted in a solution containing 1% BSA, 10% horse serum, 0.1% Tween-20, 1X PBS; subsequently, roots were incubated overnight at 4 ºC. Next day, 3 washes with 3% BSA were applied, and the secondary antibody (α-rabbit AF488 for α-RFP, and α-mouse AF488 for α-DsRed/α-mCherry) was diluted 1:500 in 1% BSA, 10% horse serum,

0.1% Tween-20, 1X PBS solution, and incubated for 1 hour at room temperature. Three washing steps with 3% BSA /1X PBS was performed.

EdU was detected by using the Click-iT EdU Alexa fluor 647 imaging kit (Thermo Fisher Scientific) according to the manufacturer's recommendations for 30 minutes. Then, roots were incubated with DAPI (10 µg/ mL) for 20 minutes, and then washed once with 1X PBS and once with water. Slides were mounted in Mowiol 4-88 (Sigma) and kept at 4 ºC until confocal microscopy observation.

3.2.2.6 MM2d cell-cycle synchronization

In 4 days old MM2d cell culture, cell cycle was arrested by elimination of sucrose from culture media during 24 h (Menges and Murray, 2002). At that point cells were synchronized in G1. Incubation proceeded in normal growth conditions and samples were taken according to cell cycle progression.

3.2.2.7 Fluorescent activated cell sorting (FACS)

Three weeks-old rosette leaves (around 30 leaves of L1 to L4) were placed in 500 µL of Galbraith Buffer (20 mM MOPS, pH 7.0, 45 mM $MgCl_2$, 30 mM sodium citrate, 0.3% Triton X-100) supplemented with 20 µg/mL of RNase A (Galbraith, 2009). Leaves were chopped with a razor blade and filtered trough a 30 µm nylon filter. Propidium iodide was added to a final concentration of 20 µg/mL and nuclei were kept on ice. Fluorescent activated cell sorting (FACS) of Arabidopsis root nuclei was performed using a FACSAria Fusion (BD Biosciciences) equipped with a 488 nm laser with 675/20-nm filter to distinguished between duplets, and a 561 nm laser with a 582/15-nm filter to distinguish between 2C, 4C, 8C and 16C cell populations. Samples were sorted at a pressure of 45 psi with an 85 µm nozzle tip. The flow data rate was adjusted to 1.2, and typically between 250–350 events/s were scored depending on the sample concentration. Nuclei were directly sorted onto 10X Nuclei Lysis buffer (Villar and Kohler, 2010), and frozen in liquid nitrogen until DNA extraction.

3.2.2.9 Fluorescent in situ hybridization (FISH)

Probe labeling and FISH were performed as described in Pavlistova et al. (2016). Briefly, root tips were hybridized in 50% formamide (Sigma)/10% dextran sulfate/2 X SSC solution, with denaturation for 2 minutes at 80 °C, and an overnight incubation at 37 ºC. Then, three washing steps with 35% formamide/2X SSC solution at 42° C for five minutes were applied, followed by two washes for five minutes in 2 X SSC at 42 °C. For the

visualization of rDNA loci the BAC clone T15P10 (GenBank AL095897/8) was used. Slides were stained with DAPI in Vectashield (Vector Laboratories) and imaged by confocal microscopy.

3.2.2.10 Terminal restriction fragment analysis (TRF)

TRF analysis was performed as in Pavlistova *et al*. (2016). Briefly, five hundred nanograms of gDNA were digested with 10 U of MseI (NEB) and processed according to Ruckova *et al*. (2008). Subsequently, TRF products were fractionated on an agarose gel, alkali-blotted onto a Hybond XL membrane (GE Healthcare Life Sciences, Waukesha, WI, USA) and detected using telomeric TR4C end-labeled with [γ-^{32}P]ATP using polynucleotide kinase (NEB). Telomeric signals on membranes were visualized using a FLA7000 imager (Fujifilm).

3.2.3 Microscopy and image analysis

3.2.3.9 Optic microscope

Seedlings and flowers were observed under a macroscope with a digital camera and external lights (LEICA). Roots, ovules, and embryos were observed in an Axioscop2 plus microscope with a CCD camera (Zeiss) with different objectives (5X/0.5, 10X/0.30, 20X/0.5, 40X/1.30 and 63X/1.4) applying DICs optic contrasts.

3.2.3.10 Confocal microscope

Confocal analyses were carried out in different confocal microscopes described below. The **Table 3.10** summarizes the conditions used in various experiments.

(i) A1R+ confocal with an Eclipse Ti-E inverted microscope (Nikon), equipped with different lasers lines (405 nm, 445 nm, 488 nm, 514 nm and 640 nm), fluorescent filters (DAPI, CFP, GFP and TRITC) and objectives (oil 20X/0.75, oil 40X/1.3, oil 60X/1.4).

(ii) LSM710 Confocal laser scanning microscope with an inverted AxioObserver microscope (Zeiss), equipped with different laser lines (Diode 405 nm, Argon 458/488/514 nm, DPSS 561 nm and HeNe 633 nm), fluorescent filters (GFP, CFP/DsRed), Nomarski filter, and different objectives (oil 25X/0.8, oil 40X/1.30, oil 63X/1.4).

(iii) LSM510 confocal with a vertical Axio Imager.Z1 M microscope (Zeiss), quipped with different laser lines (Argon 488/477/488/514 nm, HeNe 543 nm and HeNe 633 nm), fluorescent filters (FITC/EGFP and Texas-Red), Nomarski filter, and objectives (oil 25X/0.8, oil 40X/1.3).

Table 3.10. Confocal microscopes and experiments

Confocal	Line	Experiment
A1R+ Nikon	SDG7-CFP	Fluorescent Tag selection Roots stained with PI (subcelular localization) Roots stained with DAPI (subcelular localization)
	SDG24.1-mRFP	Fluorescent Tag selection Roots stained with DAPI (subcelular localization) Roots with EdU labeling Roots with IHC + EdU labeling
LSM710 Zeiss	SDG7-mRFP	Fluorescent Tag selection Roots stained with DAPI (subcelular localization)
	SDG24.1-mRFP	Fluorescent Tag selection Ovules stained with SR2200
LSM510 Zeiss	WT, *sdg7 and sdg24* mutants	Roots stained with PI for meristem length measure, WOX5 localization

3.2.3.11 Image analysis.

All Images derived from optical and confocal microscopy were analyzed with Fiji (ImageJ) and treated with Photoshop (Adobe) software.

3.2.3.11.1 Meristem measurement

The end of the meristem was determined in epidermis and cortex cell layers. In the epidermis, cell length was measured from the first cell in focus in the epidermis plane near the root tip along the file, considering the end of the meristem the first rapidly elongated cell (Casamitjana-Martínez *et al.*, 2003), that is, a cell which size doubles the average size of the cells in the meristem that will not divide again. For the case of cortex cells layer, first cell counted was the following to cortex initial cell (next to the QC) until the first rapidly elongated cell.

3.2.3.11.2 Fluorescent intensity measurement in nuclei along the root

Several images in z were acquired along the root without changing the acquisition parameters (e.g. pinhole or gain) and avoiding saturated images. The relative fluorescent intensity of each nucleus (containing DAPI and mRFP channels) was measured in Fiji by drawing circular ROIs around nuclei. DNA content was associated as the amount of DAPI

fluorescent intensity multiplied by 2 for each round of endocycle, the 2C being the lowest fluorescent intensity and 16C the highest.

Results

4.1 Candidates for Arabidopsis new histone methyltransferase

Histone modifications influence DNA replication, gene expression and the active/inactive states of the chromatin that is inherited along cell generations (Sequeira-Mendes and Gutierrez, 2015). Interpreting the histone code, referred to as the possible combinations of histone modifications, has become an interesting field of study, specially in plants, as they have evolved to retain totipotency and de-differentiation abilities. This suggests that different mechanisms for establishment and maintenance of epigenetic information diverged at some point from animals. One interesting example between divergent interpretations in the histone code is related to different methylation states of histone H4 at the lysine 20 (H4K20). In mammals, the presence of monomethylation of H4K20 (H4K20me1) has been associated to euchromatic regions and, the trimethylated H4K20 (H4K20me3) to heterochromatic areas (Schotta *et al.*, 2004; Congdon *et al.*, 2010). However, in Arabidopsis the presence of any form of H4K20 methylation has been questioned (Zhang *et al.*, 2007; Johnson *et al.*, 2008), although immunofluorescence experiments indicate that H4K20me1 is associated with chromocenters whereas H4K20me3 with euchromatin (Naumann *et al.*, 2005; Roudier *et al.*, 2011).

In animal cells the histone methyltransferases (HMTases) responsible for H4K20 methylation are PR-SET7 and SUV420H2 (H4K20me1 and H4K20meme2/3 HMTases, respectively) and contain a SET domain that confers the catalytic activity. When mutated, the ratios of H4K20me1 to H4K20me3 marks diminished, affecting directly to chromatin and cell cycle stability (Schotta *et al.*, 2004; Schotta *et al.*, 2008; Oda *et al.*, 2009; Tsang *et al.*, 2010). In Arabidopsis, the SUVH2 protein was initially described to be the putative HMTase for the H4K20me1 mark, as a decreased in this methylation mark was observed in mutant lines (Naumann *et al.*, 2005). Nonetheless, later studies did not find any functional HMTase activity associated to it (Johnson *et al.*, 2008; Kuhlmann and Mette, 2012). In this scenario, further functional analyses are needed in order to unravel H4K20 effector proteins. Here, in this study we undertook the search for the H4K20 effector proteins among the possible HMTases encoded in Arabidopsis genome.

Taking into account the opposite association of H4K20 marks in animal cells and Arabidopsis, we used mouse H4K20me3 HMTases, SUV4-20H1 and SUV4-20H2 (Schotta et al. 2004) as templates for a search of putative Arabidopsis H4K20me3 HMTases. Results from a protein BLAST provided a list of candidates, with several high-scored SET-domain sequences, including Arabidopsis SDG7/ASHH3 (ASH1 HOMOLOG-3, AT2G44150) and SDG24/ASHH4 (AT3G59960), both proteins belonging to the Arabidopsis ASH1 group (see **Table 1** in the Introduction).

As a particular detail, the naming of Arabidopsis SET domains proteins does not follow a preferential nomenclature dictation among their members; hereinafter, we will refer to our candidates as SDG7 and SDG24, respectively.

In the **Figure 4.1 A** a general alignment view shows mammalian SUV4-20H1/H2 proteins and the Arabidopsis SDG7 and SDG24 sequences aligned precisely at their SET domain-coding region. In a more detailed alignment, the amino acid similarities and identities with SUV4-20H1/H2 or between SDG7/SDG24 sequences are highlighted (**Fig. 4.1 B**). We found a 68% homology between SDG7 and SDG24 sequences, suggesting a possible redundant function among them. As SET domains have been strongly associated with lysine methylation activity, we decided to focus in the study of SDG7 and SDG24.

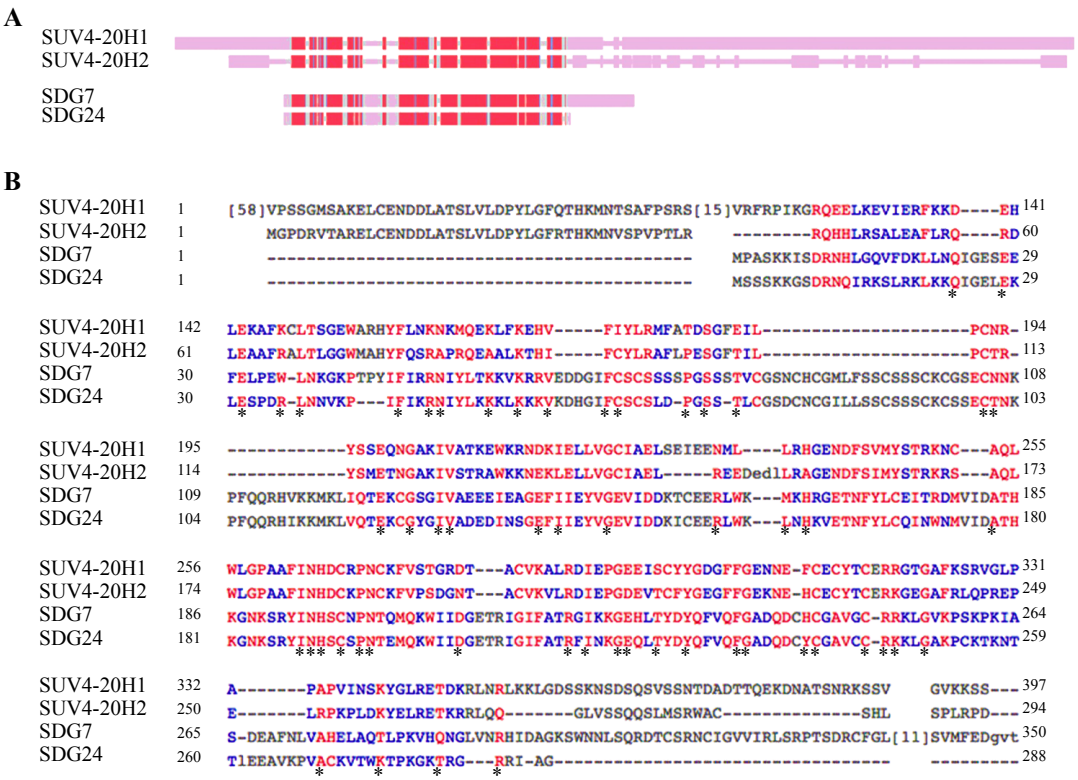


Figure 4.1. SET domain protein candidates alignment. (**A**) Schematic representation of protein domain alignment of mouse H4K20me3 HMTs SUV4-20H1/SUV4-20H2, and Arabidopsis SDG7/SDG24 proteins. In red, SET domain-coding sequences. (**B**) Detailed SET domain amino acid alignment of the proteins described in **A**. Coloured letters denote highly conserved (red) or less conserved (blue) amino acid residues. Asterisks point to equal identities with at least one of the SUV4-20H1/H2 sequence. Numbers at the edges correspond to amino acid location inside each protein.

4.1.1 Gene structure and expression analysis of candidate histone methyltransferase

4.1.1.1 *SDG7* gene structure and expression

First, we evaluated the available information for *SDG7* sequences in the Arabidopsis Information Resource (TAIR) and confirmed them by generating our own transcriptional data. The *SDG7* coding region is predicted to be 2141 bp in length, containing 11 introns and 12 exons (**Fig. 4.2 A**). By cloning *SDG7* from cDNA of wild type (Col-0) seedlings, and after sequencing, alignment between *SDG7* genomic sequence and our cDNA-transcript data confirmed the proposed gene structure described above. Next, we were interested in measuring gene expression in Arabidopsis tissues. Transcriptional maps obtained with the Arabidopsis eFP browser (Winter *et al.*, 2007) pointed *SDG7* expression in the shoot apical meristem (SAM) and during flower and embryo development with a strong signal, in rosette and caulinar leaves with mild signal, and in the RAM with a weak signal. To confirm this data, we performed an mRNA expression analysis on cDNA obtained from different tissues. qPCR results showed that *SDG7* expression is detected in roots, seedlings, leaves, flowers, and siliques at various levels (**Fig. 4.2 B**). Interestingly, relative expression among tissues revealed that there is a major peak of expression in mature flowers (when pollination already took place), followed by caulinar leaves and roots, whereas rosette leaves showed the lowest expression levels.

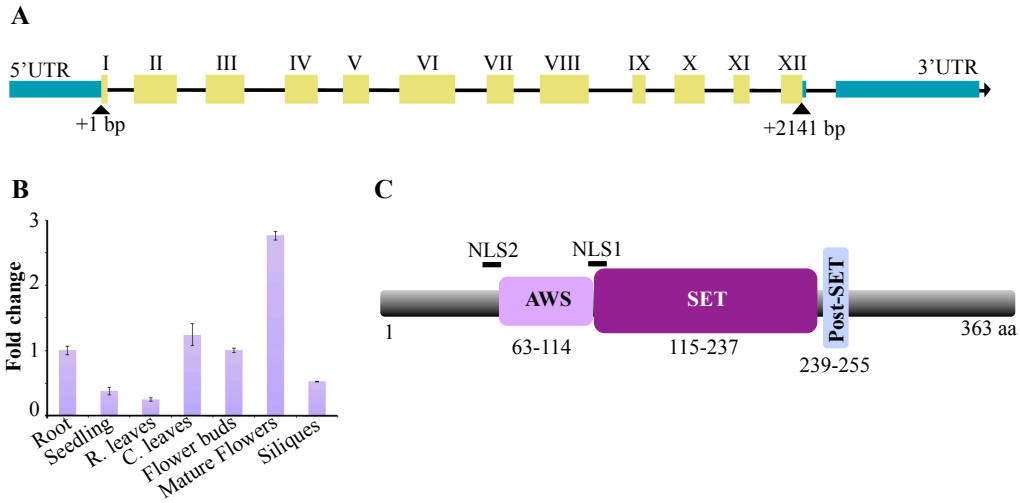


Figure 4.2. SDG7 genomic structure. (**A**) *SDG7* gene structure. Exons (yellow boxes) are arranged with roman numbers; introns as a black line; and UTR regions (5' and 3') in green boxes. Black arrowheads denote the ATG to Stop codon, confirmed by cDNA transcript analysis. (**B**) *SDG7* gene expression. qPCR from wild-type (Col-0) cDNA of 7 dps roots and seedlings, mature rosette leaves (R. leaves), caulinar leaves (C. leaves), flower buds, mature flowers, and un-dried siliques. Fold change is given by the $2^{-\Delta\Delta CT}$ method. *SDG7* amplified region is located between exons III and IV. ± SD between technical replicas. At least 2 biological replicas were performed with similar results. *ACT2* served as gene of reference. (**C**) SDG7 protein structure. Domains according to Prosite database: AWS (63-114 aa), SET (115-237 aa) and Post-SET (239-255 aa). Most probable nuclear localizations signals according to cNLS mapper are depicted as NLS1 (113-121 aa) and NLS2 (55-64 aa). Full length protein size= 363 aa.

We also analyzed *in silico* the SDG7 protein domain organization. SDG7 encodes a putative 363 amino acid protein containing a SET domain, encompassed by 2 cysteine-rich surrounding regions: associated with SET (AWS) and post-SET domains. To sustain a possible histone methylation activity, we searched for putative nuclear localization signals (NLS), and 2 sequences were found within the N-terminus of SDG7 protein sequence (**Fig. 4.2 C**), reinforcing the idea of SDG7 could be a nuclear protein.

4.1.1.2 *SDG24* gene structure and expression

When this project started, the predicted coding sequence of *SDG24* in TAIR was organized in 11 exons and 10 introns, with similar structure to *SDG7* (**Fig. 4.3 A**). However, the existing transcriptional data only offered the identification of a small transcript fragment (from the SET domain) in ovules, resulting in vague information regarding where *SDG24* could be actually expressed. Therefore, to corroborate the proposed coding sequence, we first carried out cloning attempts from seedlings, but failed. However, when we tested cDNA obtained from flowers, PCR amplification resulted in several bands (not shown) which led us to interpret the results in two ways: (i) *SDG24* was miss-annotated in TAIR, and (ii) *SDG24* could have splicing variants. After several rounds of cloning and sequencing the products derived from flowers we classified them into four alternative variants differing in their 3′ region, named: *SDG24.1*, *SDG24.2*, *SDG24.3* and *SDG24.4*, the first two being the more abundant ones, and the last two, more rare (**Fig. 4.3 A**).

In all the *SDG24* variants identified, the predicted exons X and XI were not retained in the cDNA. Our transcript analyses led us to classify *SDG24.1* variant into a 1560 bp coding sequence, containing 8 introns and 9 exons. In the case of *SDG24.2,* the ORF is predicted to be 1877 bp long, with 10 introns and 11 exons, of which exons called here IX^B and X^A were located at the 3′ untranslated exons of *SDG24.1*. With less abundant variant *SDG24.3*, an omission of part of exon VI (VI^A, VI^B), results in a shorter ORF of 1147 bp long, and a 3′ UTR sequence resembled to *SDG24.1* version. In the case of *SDG24.4*, a deletion of part of exon VIII ($VIII^A$) results in an ORF of 1390 bp and a 3′ UTR region similar to *SDG24.2*.

Although, these *SDG24* variants were found in flowers, we performed an expression analysis in different tissues of the initial part of ORF, common to all splicing variants. *SDG24* transcripts were present in roots, seedlings, leaves, flowers, and siliques (**Fig. 4.3 B**). Interestingly, the relative expression in flowers and siliques was higher than in other organs, in particular in roots, perhaps indicating that at least in roots a very low amount of transcript is present or is tissue or variant specific.

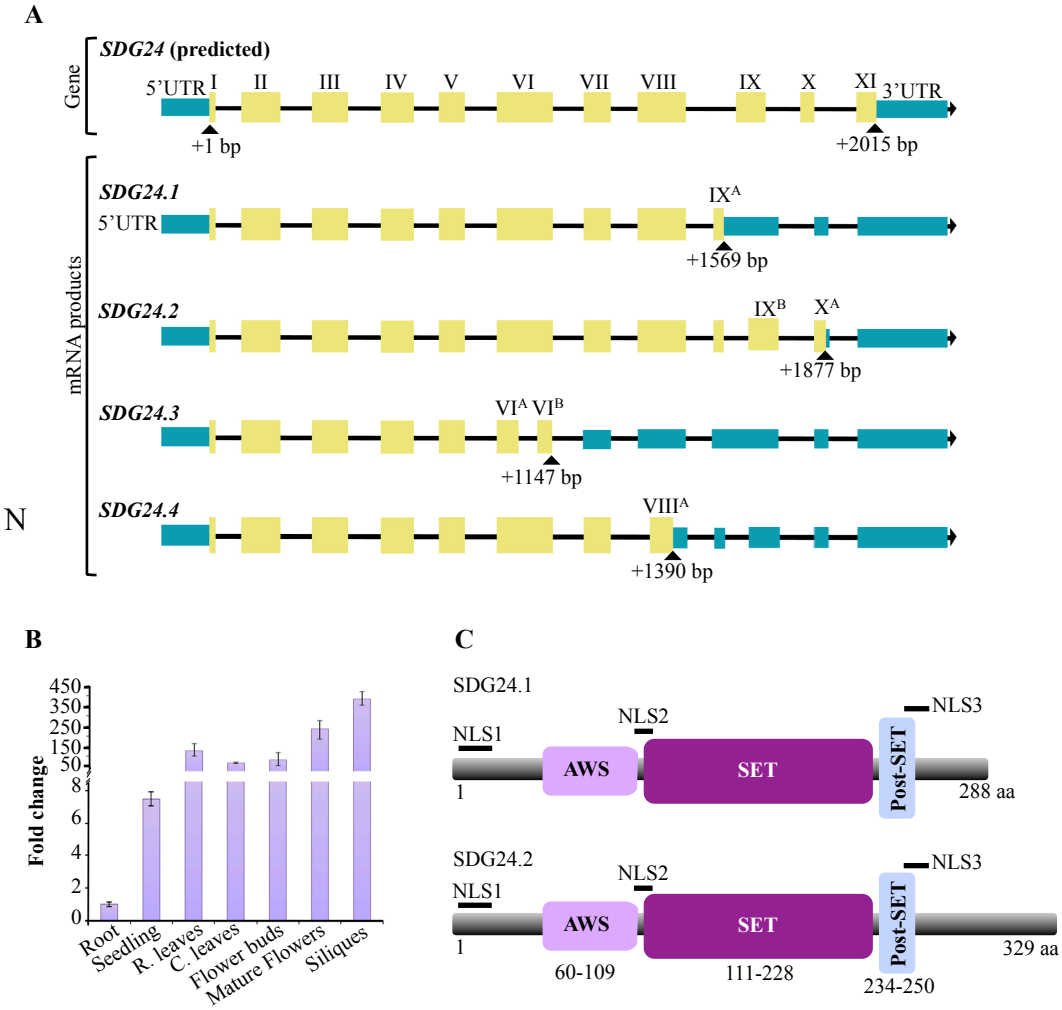


Figure 4.3. *SDG24* splicing variants structure. (**A**) *SDG24* gene structure. Exons (yellow boxes) are arranged with roman numbers; introns as a black line; and UTR regions (5' and 3') in green boxes. Arrowheads denote the ATG to Stop codons, confirmed by cDNA transcript analysis. (**B**) *SDG24* gene expression. qPCR from wild-type (Col-0) cDNA of 7 dps roots and seedlings, mature rosette leaves (R. leaves), caulinar leaves (C. leaves), flower buds, mature flowers, and un-dried siliques. Fold change is given by the $2^{-\Delta\Delta CT}$ method. *SDG24* amplified region is located between exons I and II. ± SD between technical replicas. At least 2 biological replicas were performed with similar results. *ACT2* served as gene of reference. (**C**) SDG24.1 and SDG24.2 protein structure. Protein domains according to Prosite database: AWS (60-109 aa), SET (111-228 aa) and Post-SET (234-250 aa). Most probable nuclear localizations signals according to cNLS mapper are depicted as NLS1 (5-22 aa), NLS2 (108-116 aa) and NLS 3 (246-257 aa). Domains and NLS are common for both splicing variants. Protein sizes: SDG24.1= 288 aa, SDG24.2= 329 aa.

Next, we analysed *in silico* the protein organization for the sequences of *SDG24.1* and *SDG24.2*, on which we have focused our further studies. Both of them encode two protein variants differing at their C-terminal region: SDG24.1 with 288 amino acids, and SDG24.2, with 329 amino acids. For these two variants, SET, AWS and post-SET domains are predicted to be located at the same positions. Additionally, we searched for putative NLS, and found three signals distributed along the protein sequences, NLS1 being the one possessing the highest score probability (**Fig. 4.3 C**). On the other hand, analysis in the less abundant splicing variant SDG24.3 (165 aa) the SET and Post-SET domains are altered as a direct

consequence of a shorter transcript sequence; and for SDG24.4 (241 aa), part of the post-SET domain is missing (not shown).

All this data suggest that several SDG24 variants coexist at the same time, been SDG24.1 and SDG24.2 being good candidates to posses HMT activity and nuclear localization. Additionally, the 3′ region could have a potential regulatory role as all observed variations were localized at this area, becoming more necessary an *in vivo* protein localization to discern between variants.

4.1.2 SDG7 and SDG24 protein expression analysis during Arabidopsis organogenesis

In plants, organogenesis is a post-embryonic process that occurs in a continuous manner throughout the entire lifespan. Aerial organs (leaves, stems, flowers) and underground tissues (roots, lateral roots) will be formed in order to complete the plant life cycle (Perianez-Rodriguez *et al.*, 2014). To determine the expression pattern of our candidate proteins, we generated transgenic plants expressing tagged versions under the control of their endogenous promoters (**Fig. 4.4**). For SDG7 and two SDG24 variants, SDG24.1 and SDG24.2, three tagged versions were generated: β-glucuronidase protein (GUS), enhanced cyan fluorescent protein (CFP) and the monomeric red fluorescent protein (mRFP). These reporters were chosen to facilitate their use with others, e.g. GFP.

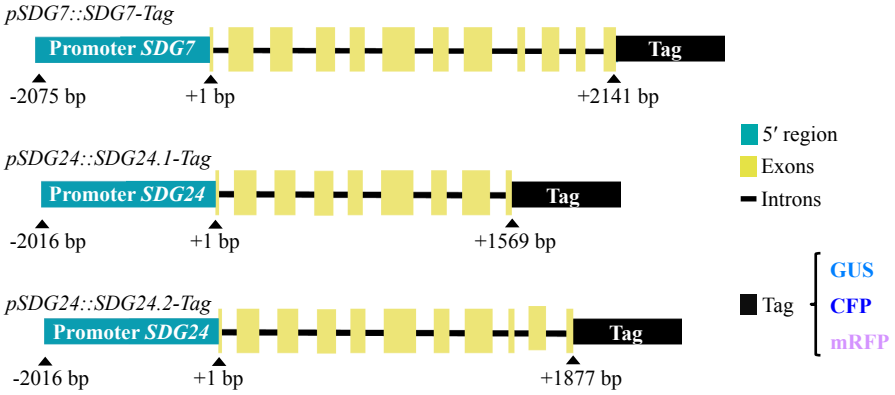


Figure 4.4. *SDG7* and *SDG24* tagged lines. The genomic sequences of *SDG7* and *SDG24* were cloned including ~2 kb of their 5′ region and fussed to different tags (GUS, CFP and mRFP). Two versions of *SDG24* were generated, *SDG24.1* and *SDG24.2*, differing in their 3′ region according to the mRNA analysis. The schematic representation shows exons (yellow boxes) and introns (black line) of the genomic sequence. The arrowheads denote positions in the genomic sequences before or after (-, +) the ATG starting codon.

As it could be expected for the *SDG24* gene that possesses different spliced variants, an *in silico* analysis on the *SDG24.1* and *SDG24.2* genomic constructs was performed to rule out

the production of alternatively tagged products (not shown). The outcomes from those constructions will only produce the *SDG24.1* and *SDG24.2* specific variants that are in frame with the different tags; however, we cannot discard the occurrence of other(s) uncharacterized splicing maintaining in frame the tag expression.

4.1.2.1 SDG7 is expressed in leaves, roots, flowers and embryos

We studied the expression of SDG7 in different Arabidopsis developmental stages. The pSDG7::SDG7-GUS line allowed us to located it in the root, lateral roots, lateral root primordia (LRP), shoot apical meristem (SAM), rosette leaves, inflorescent meristem (IM) (**Fig. 4.5**), and during flower and embryo development (**Fig. 4.6**). Interestingly, we noticed that SDG7 expression had an ON-OFF pattern in all observed tissues, being higher in proliferative cells, and diminished progressively when differentiated growth took place.

The ON/OFF signal was visualized during aerial organ development starting with the formation of rosette leaves. Protein accumulates covering the total area of the first pair of leaves from day 2 to day 6; then, while cell expansion progresses, SDG7 is confined to the basal part of each leaf between days 7 to 9, to finally disappear at day 10 (**Fig. 4.5 A-L**). This pattern was repeated for all emerging pair of leaves, and continued when floral transition aroused. At this stage SDG7 is detected only in the last emerged leaves and in the new IM in a rosette formation (**Fig. 4.5 M**).

Incubation of SDG7-GUS roots for 30 minutes in GUS staining solution showed a particular high-accumulated signal at the meristem zone that decreased at elongation and differentiation zones, until no signal was detectable (**Fig. 4.5 N**, **O**). Additionally, SDG7 expression was observed during lateral root formation, being present in all developmental stages and maintained in the meristematic zone, as lateral roots were elongated (**Fig 4.5 N**, **P**).

After the flowering transition, SDG7 is highly expressed in all flower buds between floral stages 1 to 6 (**Fig. 4.6 A**). Afterwards, we visualized the expression in growing petals and carpel development between floral stages 7 to 11. Interestingly at flower stage 11, meiosis of the megaspore takes place (major events of male and female gametophytes development are summarized in **Fig. 1.9** in the Introduction). Curiously, between flower stages 12 and 13, SDG7 signal was undetectable. After pollination, SDG7 protein became more apparent inside the carpel, letting us to investigate its expression during embryo development. In the growing zygote, SDG7 also showed the ON/OFF pattern, where signal was accumulated in proliferative cells from early embryo stages to early torpedo (**Fig. 4.6 B1-7**).

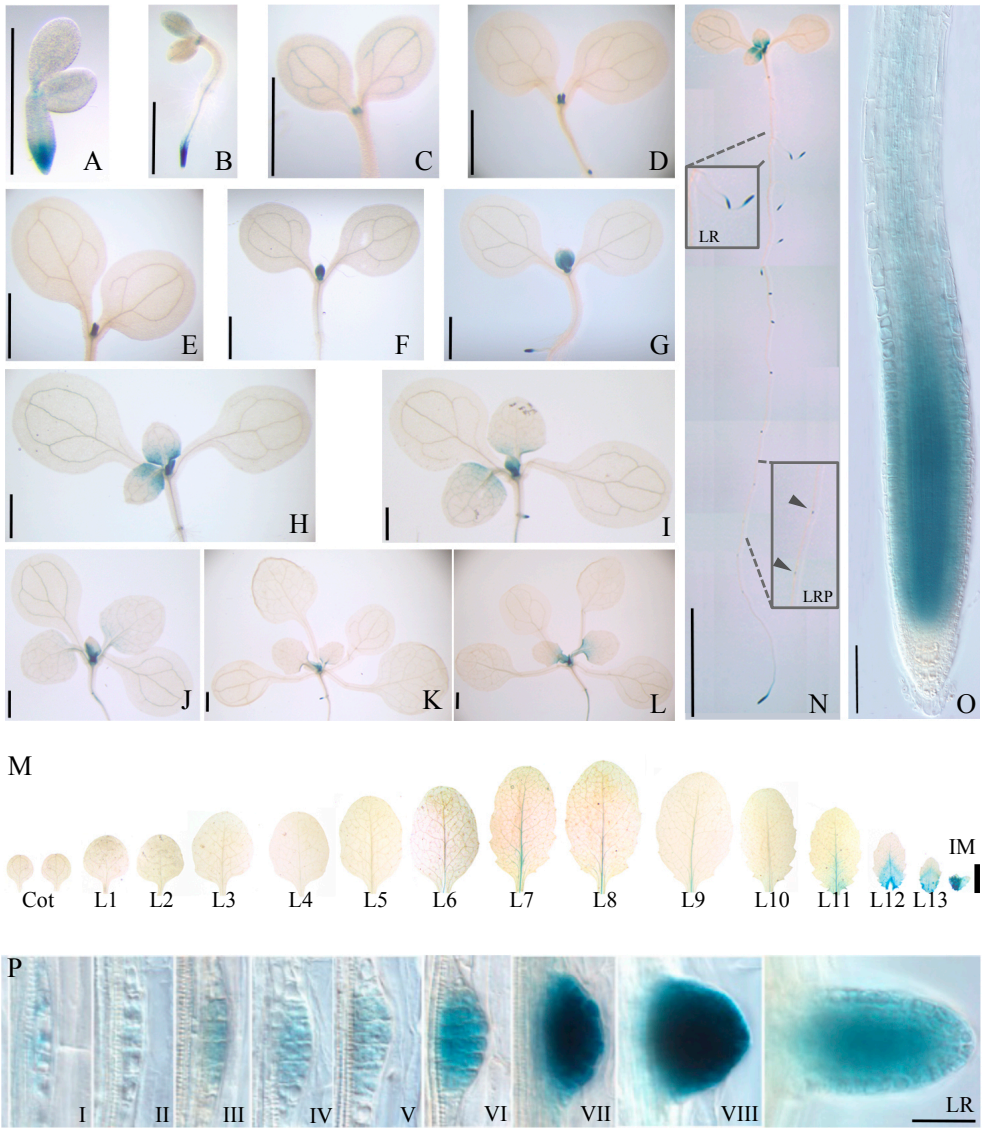


Figure 4.5. SDG7 is expressed in leaves and roots. Expression of SDG7-GUS in Col-0 background. (**A-L)** Time course expression in seedlings, in order from 1 dps (**A**) to 12 dps (**L**). Bars=1 mm. (**M**) Flowering transition in a 21 dps plant. Cotyledons (Cot), rosettes leaves (L1-L13) and inflorescent meristem (IM) are depicted (tissues come from the same rosette). Bar=3 mm. (**N**) 8 dps seedling showing lateral roots (LR, upper inside box) and lateral root primordia (LRP, lower inside box). Bar= 5 mm. (**O**) Optical image of primary root at 6 dps. Bar= 100 µm. (**P**) Optical images of LRP formation stages (I-VIII) and a fully emerged LR. Bar= 50 µm.

Also, we detected SDG7 signal at the chalazal endosperm, a pocket of cells with a low rate of cellularization (Berger, 1999). After cotyledons and radicle (the embryo root) start to elongate, SDG7 expression is confined at the meristem reservoir of SAM and RAM (**Fig. 4.6 B8-9**). Detailed events of embryo development are summarized in **Fig. 1.10** in the Introduction.

Taken together all *in vivo* expression data we conclude that SDG7 protein is highly expressed in proliferative growing tissues, and its expression progressively decreased while differentiation occurs in the formation of organs.

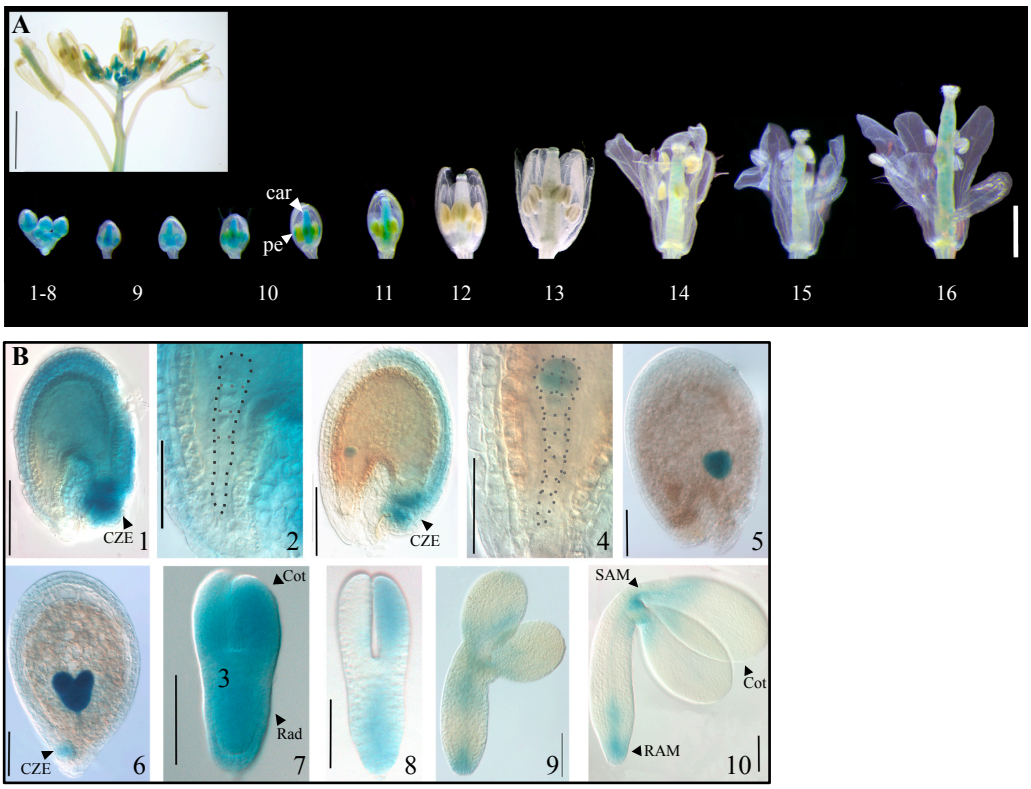


Figure 4.6. SDG7 is expressed during flower and embryo development. (**A**) Detection of SDG7-GUS in flower stages 1-16. Bar= 1 mm. A fully developed inflorescence is shown inside the box. (**B**) Optical images of embryos at different stages: 4-cell zygote (1-2), octant (3-4), early heart (5), late heart (6), torpedo (7, 8), walking-stick (9) and mature embryo (10). B2 and B4 are amplifications of B1 and B3 respectively. car: carpel; pe: petal; CZE: chalazal endosperm; SAM: shoot apical meristem; RAM: root apical meristem; Cot: cotyledons; Rad: radicle. Bars= 100 µm; and 50 µm in 2B and 4B.

4.1.2.2 SDG24 is expressed in vegetative organs

We followed SDG24 expression pattern with the aid of GUS reporter lines. First, we decided to find out if the SDG24.1-GUS variant was expressed in the root, as initial transcriptional data did not discerned between variants. Short incubation times with GUS staining solution (between 30 minutes up to 8 hours) did not showed any positive signal. Surprisingly, after prolonged incubation periods (>24 hours) a signal was successfully detected in the root (**Fig. 4.7 B**). Protein also accumulated in LRPs and in fully emerged lateral roots (**Fig. 4.7 A**, **F**). Intriguingly, we also detected SDG24.1 expression in nuclei from young and mature trichomes from 10-12 dps rosettes leaves (**Fig. 4.7 C-E**), although no signal was observed in rosette or caulinar leaves. We also performed the examination of the SDG24.2-GUS variant in the vegetative organs; although no signal was detected even when prolonged GUS incubation periods were applied (not shown).

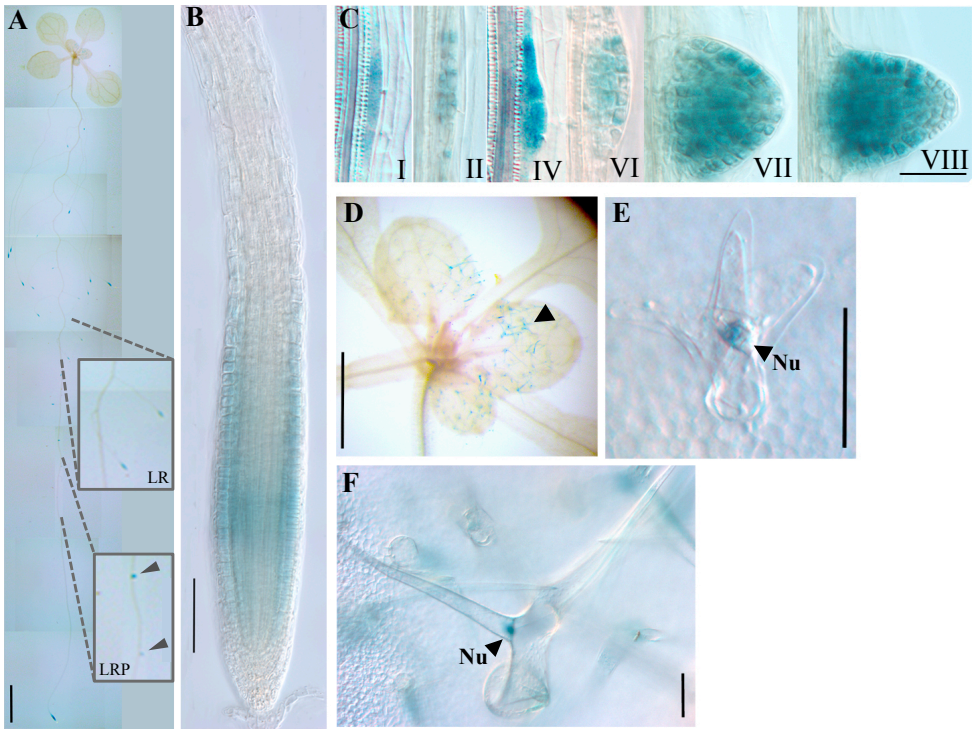


Figure 4.7. SDG24.1 is expressed in roots and trichomes. Detection of SDG24.1-GUS in Col-0 background. (**A**) 12 dps seedling showing lateral roots (LR, upper inside box) and lateral root primordia (LRP, lower inside box). Bar= 5 mm. (**B**) Optical image of primary root at 6 dps. Bar= 100 µm. (**C**) Optical images of LRP formation stages (I-VIII). Bar= 50 µm. (**D-F**) Trichomes of 12 dps rosette leaves. Optical image of young (**E**) and mature trichome (**F**) expressing SDG24.1 at the nuclei (Nu).

From these results we conclude that the SDG24.1 is expressed in vegetative organs, including the root and the nuclei from trichomes, while the SDG24.2 variant is absent in those tissues.

4.1.2.3 SDG24 is expressed in reproductive organs

Since the *SDG24* transcripts were highly enriched in flowers, we followed *SDG24.1* expression during flower development (**Fig. 4.8 A**). In particular, the SDG24.1-GUS signal was only detected after prolonged incubation periods (> 48 hours) with a first peak of expression in the anthers between early floral stages 7-11. At this point, pollen microspores are formed and the first mitotic division takes. Curiously at stage 12, protein was undetectable in any floral organ. Then, between late-12 and 13 floral stage, when megaspores reaches their maturity, signal became very high at ovules inside the carpel and fainted after pollination took place (**Fig. 4.8 A**, **B**).

To confirm that SDG24.1 expression in the embryo sac occurs in a very particular stage without pollination, we emasculated 12-stage and 13-stage flowers and let them grow under

normal conditions over 18 hours; after GUS detection, the signal was visualized only in the new 13-stage flower and absent in the un-pollinated 14-stage ones (not shown).

After that, we performed a more detailed analysis in ovules from 13-stage flowers. The mature embryo sac consists of 8-nucleated, 7-cells organ: the central cell (1 cell with 2 nuclei) and egg cell (1) being the female gametes; and the synergids (2) and antipodals (3), the accessory cells the (Bowman, 1994; Schneitz *et al.*, 1995; Baroux and Grossniklaus, 2015). To analyze protein expression at the cellular level, we used the SDG24.1-mRFP reporter line. Expression was detected in the two polar nuclei of the central cell before fusion (8% out of 35 ovules analyzed), in secondary nucleus of central cell after fusion (90%), and both central cell and egg apparatus (~3%), with the egg cell and synergids being indistinguishable at the micropylar pole (**Fig. 4.8 C-E**). With a prolonged GUS enzymatic activity of SDG24.1-GUS line (48-72 hours), expression in the egg apparatus became more apparent (80% out 50 ovules examined) but inconclusive to determine the specific cell types (**Fig. 4.8 F-H**).

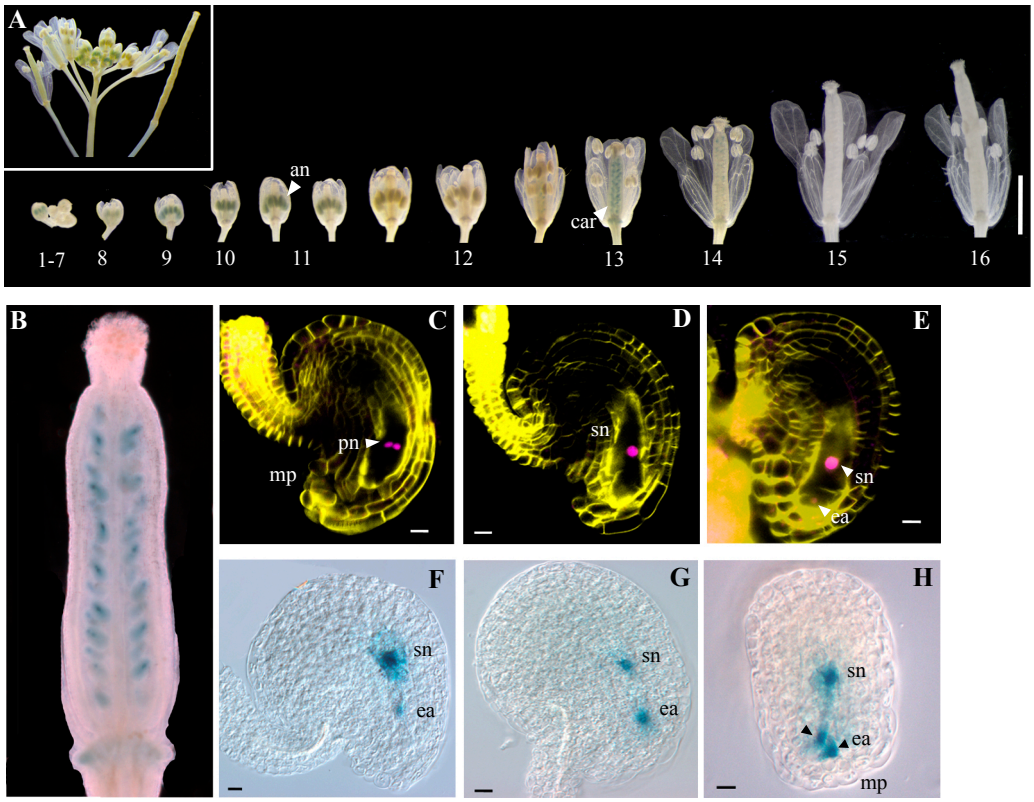


Figure 4.8. SDG24.1 is expressed during flower and embryo sac development. (**A**) Expression of SDG24.1-GUS during flower stages from 1-16. Bar= 1 mm. A fully developed inflorescence is shown inside the box. (**B**) Optical image of a 13-stage emasculated floral carpel. (**C-H**) Whole-mount ovules collected after 24 h from emasculated 12-stage flowers. Mid-plane confocal images of ovules expressing SDG24.1-mRFP (magenta) in two polar nuclei from central cell (**C**), the secondary nucleus of the central cell (**D**), and the secondary nucleus and egg apparatus (**E**). The SR2200 dye (yellow) is used as membrane counterstain. (**F-H**) Mid-plane optical images of ovules expressing SDG24.1-GUS in secondary nucleus and egg apparatus. Triangles in H point to two different stained nuclei. an: anther; car: carpel; pn: 2 polar nuclei from central cell; sn: secondary nuclei from central cell; ea: egg apparatus; mp: micropylar pole. Bars= 10 µm.

Taken together this data, we can conclude that: (i) SDG24.1 is expressed in male and female reproductive organs, (ii) the expression of SDG24.1 in ovules, as well as in trichomes, confirms the nuclear localization of SDG24.1, (iii) SDG24.1 is expressed at least in one of the female gametes, the central cell, just prior pollination, acting as a possible indicator for embryo sac maturity, and (iv) prolonged incubation periods in GUS staining solution might suggest low SDG24.1 protein expression in the tested organs.

On the other hand, we performed the examination of SDG24.2 variant with the aid of the SDG24.2-GUS reporter line. Interestingly, a positive signal was detected only in the anthers between floral stages 7-11 (**Fig. 4.9**). We also noticed that the signal decrease progressively until no signal was visualized at floral stage 13, when pollen has completed mitotic divisions and it is ready to be dehiscent. With these results, we can conclude that SDG24.2 is specifically expressed at early anther development.

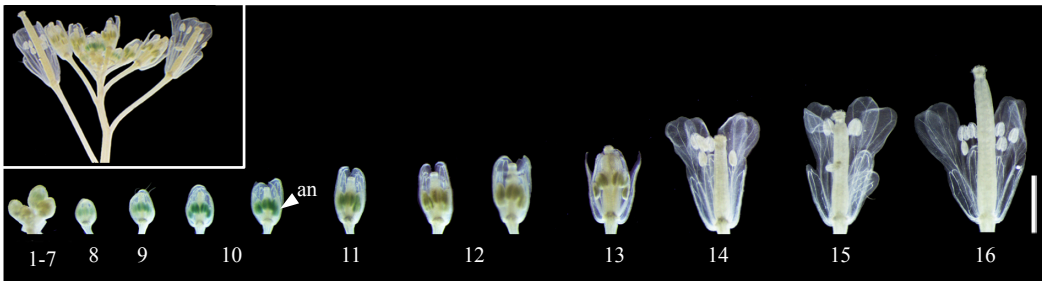


Figure 4.9. SDG24.2 is expressed during flower development. Expression of SDG24.2-GUS in Col-0 background. Flower stages from 1-16. A fully developed inflorescence is shown inside the box. Bar= 1 mm.

4.1.2.4 SDG7 has a nuclear expression

A prerequisite to function as a chromatin modulator is to possess nuclear localization. Using the SDG7-GUS lines was difficult to determine any particular nuclear pattern. Thus, we decided to analyze fluorescent reporters in the root meristem, as confocal microscopy could be more sensitive to detect subcellular localization. With the aide of SDG7-CFP lines, we found that SDG7 was present in the nuclei of all root meristem cells (vascular, endodermal, cortex, epidermal, lateral root cap (LRC), and the columella), being more clear in epidermal cell nuclei, and less intense in the columella nuclei (**Fig. 4.10 A**, **B**).

We also noticed a high fluorescence background around the nuclei, indicating that expression was also located at the cytoplasm. In cells outside the RAM the SDG7 expression signal was confined to the nuclei, and it was even more nuclear at the beginning of the EZ,

probably because the cytoplasm retracts with the increasing vacuole size (**Fig. 4.10 A**). As seen with SDG7-GUS lines, the signal of SDG7-CFP decreased progressively along the EZ and DZ until is no longer detected (not shown).

We also evaluated SDG7-mRFP expression lines. This construct allowed us to sharp the detection in the nuclei from the root and lateral root primordia (not shown). We were also able to colocalize SDG7-mRFP with DNA, by applying a soft DAPI (4′,6-diamidino-2-phenylindole) staining procedure (0.25 µg/mL of DAPI solution for five to ten minutes) (**Fig. 4.10 C**).

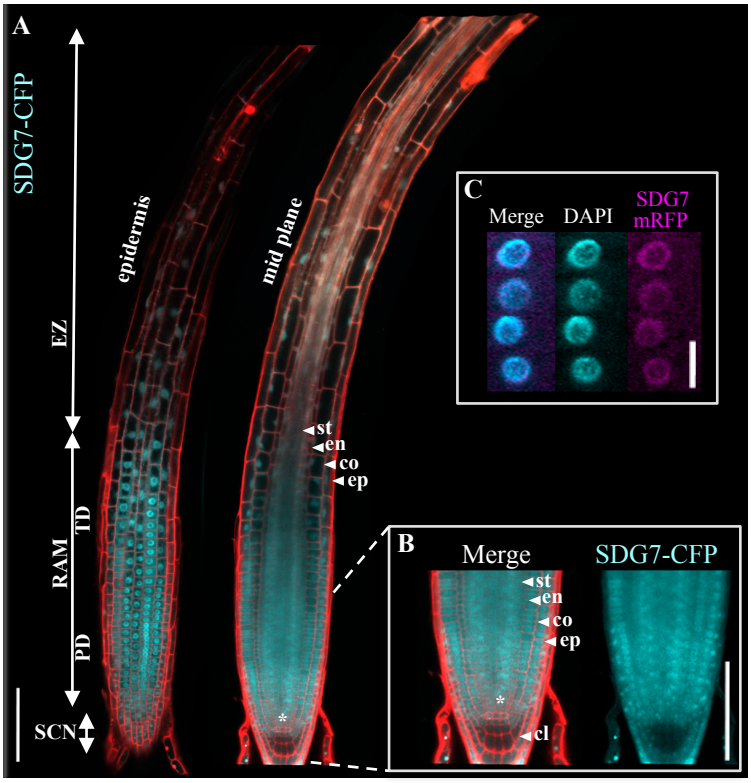


Figure 4.10. SDG7 is expressed in the nuclei and cytoplasm of RAM cells. Confocal images of roots expressing SDG7 tagged lines. (**A**) SDG7-CFP (cyan) expression in all cells layers of the RAM at 6 dps. Epidermal cell layer (right) or mid plane section (left). Bar= 100 µm. (**B**) Amplified mid plane root section of **A**. The QC is indicated by an asterisk. Bar= 50 µm. Propidium iodide (red) is used as cell wall counterstaining. (**C**) RAM nuclei from the same cell file expressing SDG7-mRFP and stained with DAPI (cyan). Bar= 10 µm. SCN: stem cell niche; QC: quiescent center; PD: proliferation domain; TD: transition domain; EZ: elongation zone; ep: epidermis; co: cortex; en: endodermis; st: stele; cl: columella.

Noteworthy, prolonged incubation periods with DAPI, highest concentrations or paraformaldehyde fixation produced a massive loss of the SDG7 fluorescent signal from the nuclei to the cytoplasm, affecting some cells or the entire cell file from all cell layers (not shown). Furthermore, an important technical feature of nuclei fluorescent detection was the need of very aggressive confocal conditions, where the laser was usually set between 80-95% capacity and low scanning rate, which at the end, resulted in fluorescent bleaching and cell damage on *in vivo* detections. Also, confocal analysis for both SDG7-CFP and mRFP tagged lines was only achieved with specific laser lines and detectors for those wavelengths, being impossible to distinguish them with more wide-range detectors in older confocal microscopes available in the CBMSO facilities. Confocal equipment and conditions are described in detail

in Materials and Methods.

From these results we conclude that SDG7 is expressed in nuclei and cytoplasm of root cells, although the signal is only detected with the use of harsh confocal settings. Also, SDG7 is prone to degradation with the application of chemical stresses, adding technical difficulties for molecular characterization.

4.1.2.5 SDG24 has a nucleolar organization in the root

After nuclear localization of SDG24.1 in trichomes and the gametic central cell, we also wondered if this pattern could be visualized at the root with the aid of fluorescent-tagged lines and confocal microscopy. As an interesting fact, we fail to observed any expression with CFP reporter, been analyzed more than 30 F2-transgenic lines presenting the typical 3:1 segregation pattern (one T-DNA insertion event). With the SDG24.1-mRFP line we found a signal in nuclei of all cell layers of the root (**Fig. 4.11 A**).

To confirm this localization, we also stained nuclear DNA by performing a rapid fixation of the roots with paraformaldehyde, followed by DAPI staining (diluted at 0.25 µg/mL) for 10 minutes. Interestingly, DAPI staining narrowed SDG24.1 localization to the nucleoli (**Fig. 4.11 B**, **C**). This nucleolar pattern is more clearly observed after 3D reconstruction, where SDG24.1 is occupying the unstained DAPI regions, as less DNA condensates in this area (**Fig. 4.11 D**). Amplification of the nucleolar area also revealed aggregation patterns or clusters where SDG24.1 was brighter (**Fig. 4.11 E**). This pattern might correspond to protein aggregation in different nucleolar compartments at that particular cell stage.

The size of nucleolus reflects the dynamics of ribosome production and change according to the cell demand for protein biosynthesis. In proliferating cells, where ribosome production is high, the nucleolus diameter can be 20 times bigger than those of differentiated cells (Ruda and Warner, 2004; Olson, 2011). In 3D-reconstructed nuclei from the same root cell file we observed this pattern, where DAPI stained nuclei at the TD of the RAM possess larger unstained areas than more distant nuclei in the EZ (**Fig. 4.12 B**). Accordingly, SDG24.1-mRFP expression was restricted to the nucleolar area of each nucleus (**Fig. 4.12 A**, **B**). Curiously, we also observed that nucleolar domain (and SDG24.1 expression as well) was not always localized at the center of nuclei, but rather displaced towards the nuclear envelope, being more evident in cells at the EZ with smaller nucleoli (**Fig. 4.12 B**, **C**). In agreement with this observation, the nucleolar proximity or direct contact to the nuclear envelope has been described in yeast an higher eukaryotes to be associated with local

modification of nuclear membrane rigidity and peripheral tethering of rDNA (Mekhail *et al.*, 2008; Olson, 2011; Wang *et al.*, 2016).

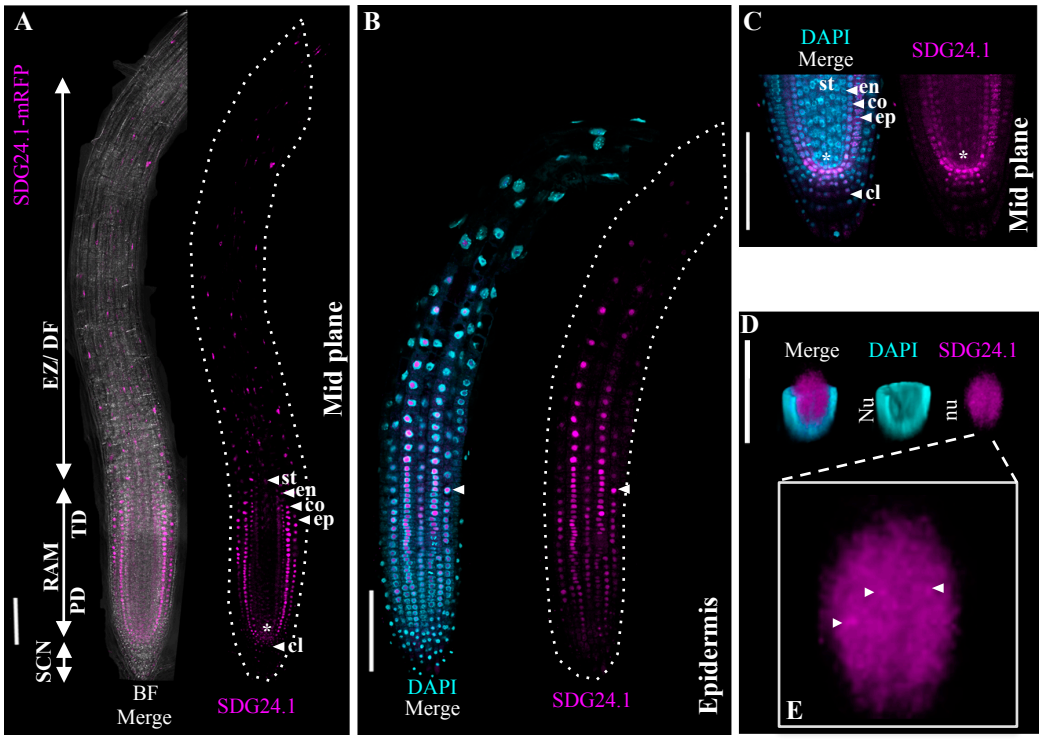


Figure 4.11. SDG24 is expressed in nucleoli of root cells. (**A-B**) Confocal images roots expressing SDG24-mRFP (magenta) in the RAM on 6 dps roots. (**A**) Mid plane section of RAM showing bright field (BR) and SDG24-mRFP expressing in all root cell layers and QC (*). (**B**) DAPI stained root (cyan) showing the epidermal cell layer. Arrowheads points to stained nuclei with SDG24 at the nucleolar area. (**C**) Same as in **B**, but with the QC plane of the RAM. Bar= 50 µm. (**D**) 3D reconstruction of a RAM nucleus (*z*-50 planes). Nu: nucleus, nu: nucleolus. Bar= 10 µm. (**E**) Amplification of nucleolus from **D**. Arrowheads detailing SDG24 brighter aggregation spots. Nuclei from B-D are counterstained with DAPI (cyan). SCN: stem cell niche; QC: quiescent center; PD: proliferation domain; TD: transition domain; EZ: elongation zone; DF: differentiation zone; ep: epidermis; co: cortex; en: endodermis; st: stele; cl: columella. Bar= 100 µm (A, B).

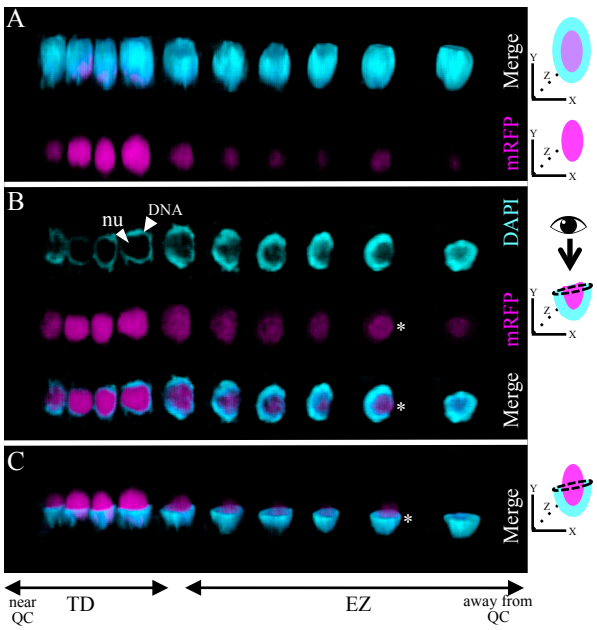


Figure 4.12. 3D reconstruction of SDG24 expressed in root cells. 6 dps root nuclei expressing SDG24.1-mRFP. A section of one epidermal cell file stained with DAPI (cyan) is presented in different orthogonal projections: front view (**A**), upper view from transversal plane (**B**), front view with a transversal cut of DAPI area leaving SDG24.1 expression uncovered inside the nucleus (**C**). Arrowheads in **B** point to stained DNA and nucleolar area (nu) from one nucleus. Asterisks (*) in B and C point to displaced nucleolar area towards the nuclear membrane. Cell positioning at the RAM is indicated at the bottom: TD: transition domain; EZ: elongation zone.

It must be emphasized that evaluations of mRFP lines were not exempt of technical difficulties. In general, a low detection signal in the root was appreciated. Also, we detected loss of fluorescence as a consequence of root handling, where mechanical manipulation of roots for confocal observation, led to degradation of SDG24.1-mRFP nucleolar pattern in several nuclei from the same cell file, and in some cases in an entire cell layer. Mechanical stress also produced autoflorescent particles surrounding the nuclear area where damage has occurred (not shown). This data suggest, that the nucleolar pattern of SDG24.1 in the root is rapidly altered when mechanical handling is applied, maybe as a stress response mechanism.

From the confocal analysis we conclude that at least in the root, SDG24.1 has a nucleolar expression in different root cell layers and particularly seems to aggregate in clusters inside the nucleolus. Additionally, SDG24 is sensitive to mechanical handling, adding technical difficulties to molecular characterization.

4.1.3 SDG7 and SDG24 expression during the cell cycle

4.1.3.1 SDG7 and SDG24 are absent during mitosis in the RAM

One of the most significant features that became evident during SDG7 and SDG24.1 protein localization in the root was their absence in mitotic cells. Mitotic cells are frequent in the PD of the RAM, and less common at the TD. Consequently, we examined mitotic figures inside the RAM with the aid of DAPI staining. Mitotic figures did not show any detectable SDG7-mRFP or SDG24.1-mRFP proteins, from prophase to anaphase (**Fig. 4.13 A**, **B**).

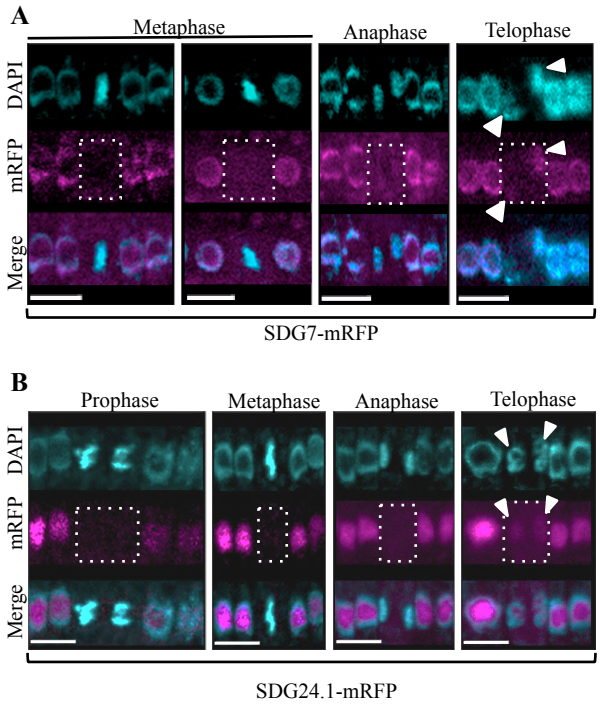


Figure 4.13. SDG7 and SDG24 are absent during mitosis. Confocal images of mitotic figures of 6 dps roots expressing SDG7-mRFP (**A**) or SDG24.1-mRFP (**B**) in magenta. DAPI was used for DNA counterstaining (cyan). White triangles in Telophase show the accumulation of new SDG7-mRFP or SDG24.1-mRFP protein. Bars= 10 µm.

To observe telophases, DAPI background allowed us to search for engrossed cell areas surrounded by smaller nuclei. Interestingly, at telophase, SDG7 and SDG24.1 proteins started to accumulate once again at the nuclear area. These results could be explained at least for the nucleolar SDG24.1 protein, as nucleoli assembly is an early event initiated at telophase (Hernandez-Verdun, 2011). In resemblance to SDG24.1, our results led us to speculate that SDG7 might have as well a nucleolar functional pattern.

4.1.3.2 SDG24 is present in G1-S-G2 phase cells in the RAM

Our next aim was to answer if our candidate proteins were stably expressed throughout the rest of the cell cycle. To do so, we followed several approaches, including the labeling with EdU (a nucleoside analog of thymidine) and the combination with fluorescent cell-cycle markers. Here, we focused on SDG24 protein. In our first attempt to determine if SDG24.1-mRFP is expressed in S-phase cells, we labeled with EdU, fixed the roots, immunodetected the mRFP-tag and stained with DAPI to highlight all nuclei (as detailed in Methods). Unfortunately, after testing different mRFP antibodies and dilutions, we were not able to detect SDG24 (**Fig. 4.14 A**). Interestingly with the control protein, the histone H3.3 (HTR5-mRFP; Otero et al., 2016), we detected a clear nuclear colocalization of the highly expressed histone, but not so efficiently for the less expressed ORC1b protein (ORC1b-mRFP; Vergara, 2017).

These results led us to investigate if the protein detection range of mRFP antibody was influencing these results. To that end, we performed a Western blot analysis of nuclear protein extracts from roots expressing SDG24.1-mRFP, and HTR5-mRFP and ORC1b-mRFP as controls. Indeed, the antibody only detected a protein signal from HTR5-mRFP up to 1.5 µg of nuclear extracts (**Fig. 4.14 B**, **C**). Although ORC1b-mRFP was not detected by its mRFP tag, it was visualized with the use of an ORC1b specific antibody, indicating that ORC1b degradation did not occur (**Fig. 4.14 C**). Therefore it seems that the amount of SDG24.1 is below the detection level of the antibodies used.

Interestingly, ORC1b is degraded by the proteasome pathway and nuclear protein extractions needed to be done in the presence of proteasome inhibitors (Vergara and Gutierrez, 2017). To rule out that a similar process was damaging SDG24 protein during nuclear extractions, we incubated roots expressing SDG24.1-mRFP in a mix of proteasome inhibitors (50 µM MG132 plus 0.5 µM epoxomycin) for 3 hours (**Fig. 4.15**).

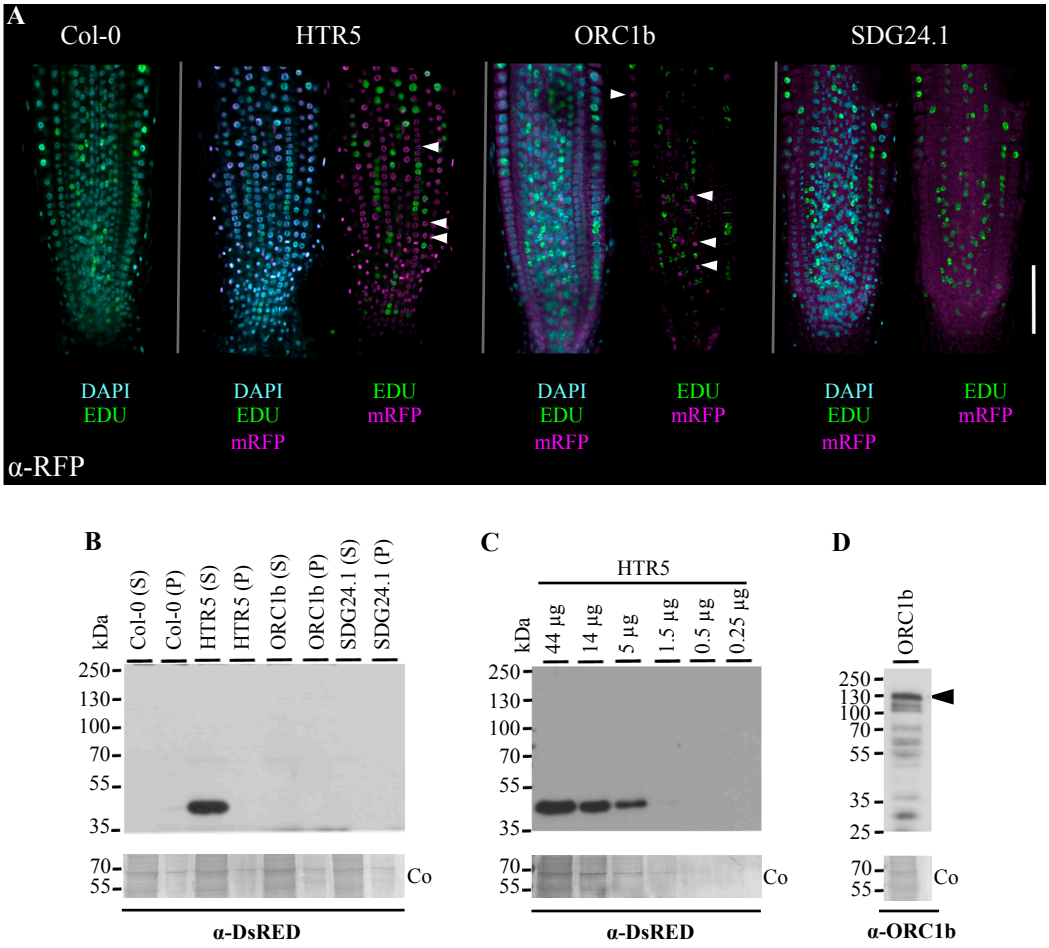


Figure 4.14. SDG24 protein can not be identified by immunocytochemistry or Western blot. Roots expressing SDG24.1-mRFP and different mRFP-Tagged proteins (as positive controls) were subjected to immunodetection with different antibodies. (**A**) Confocal images of 15 min EdU labelled 5 dps roots from Col-0 (wild-type), HTR5-mRFP, ORC1b-mRFP and SDG24.1-mRFP. In magenta, immunohistochemical detection of mRFP nuclei (α-RFP, Thermo-Fischer Scientific). In green, EdU positive nuclei detected by Click-it reaction. DAPI was used as counterstaining (Cyan). Arrowheads point to positive mRFP nuclei. Bar= 50 µm. (**B-C**) Western blot detection of mRFP-tagged proteins presented in **A**. (**B**) Nuclear-enriched protein extracts from 12 dps roots were subjected to WB analysis with α-DsRED antibody (Santa Cruz Biotechnologies). 44 µg of soluble (S) and insoluble (P) protein fractions were loaded. ORC1b was treated with proteasome inhibitors during protein extraction to avoid protein degradation. (**C**) Soluble fraction of HTR5-mRFP presented in **B** was diluted to determine the sensitivity range of α-DsRED antibody. (**D**) Equal amount of ORC1b soluble fraction presented in **B** was detected by specific α-ORC1b antibody. The arrowhead shows ORC1b-mRFP location. Molecular markers in kDa are shown at the left of each figure. Coomassie stained membranes (Co) are shown in the bottom as protein loading control. Expected protein sizes: HTR5-mRFP, 47 kDa; ORC1b-mRFP, 120 kDa; SDG24.1-mRFP, 70 kDa.

Observations at the confocal microscope revealed that ORC1b-mRFP increased its florescent pattern in the root (as expected), but no major differences were observed for SDG24.1-mRFP, suggesting that SDG24 is not targeted to proteasome degradation. Altogether these results strongly indicate that SDG24 is indeed a low expressed protein in the root, and available antibodies for mRFP tags are not efficient to detect it neither by immunohistochemical assay nor by Western blot.

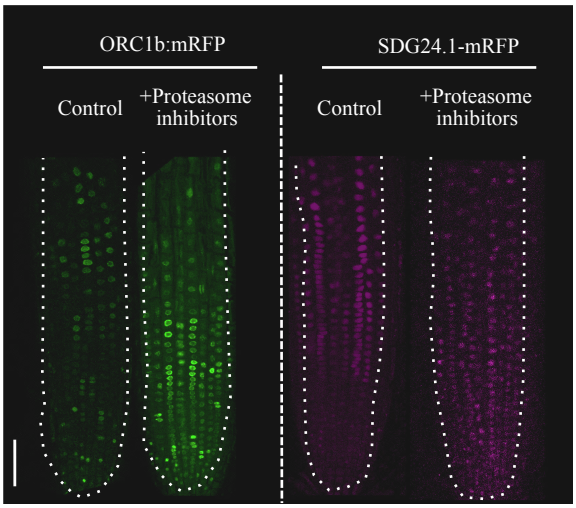


Fig. 4.15. SDG24 is maintained in the presence of proteasome inhibitors. Confocal images of 6 dps roots expressing ORC1b-mRFP (green) or SDG24.1-mRFP (magenta) treated with proteasome inhibitors (50 µM MG132 plus 0.5 µM epoxomycin) for 3 hours. Control plants are shown. Bar= 50 µm.

So far, none of the conditions tested allowed us to detect SDG24 in S-phase cells. Thus, we decided to use an *in vivo* approach, where EdU labeled cells could be colocalized with the fluorescent signal of SDG24.1-mRFP. To do so, roots were incubated with EdU for 15 minutes, fixed for 5 minutes, permeabilized for 15 minutes, and EdU detected with the Click-it reaction. After 10 minutes of DAPI staining roots were examined by confocal microscopy (see Methods for more detailed protocol). With this method, EdU labeled cells were positively detected colocalizing with SDG24.1-mRFP nucleoli along the RAM in epidermis and cortex cells files (**Fig. 4.16 A**). Fluorescence damage was observed for all treated roots. In cell files with positive SDG24 fluorescence, we localized 100% of EdU labeled cells with mRFP expression, being present in all of four (early-to-late S-phase) EdU expression patterns (**Fig. 4.16 B**) (Dvorackova *et al.*, 2017). Cells without SDG24.1 expression within the same cell file were also EdU negative (**Fig. 4.16 A**), indicating that the absence of SDG24.1 was not related to S-phase. Therefore, we conclude that SDG24 is present during the entire S-phase.

To detect SDG24.1 expression in other cell cycle phases, we crossed them with ORC1b-GFP (Vergara, 2017) and CYCB1;1-GFP (Ubeda-Tomas *et al.*, 2009) markers for identification of G1 and G2 cells, respectively. ORC1b is a subunit of the origin recognition complex (ORC) that is synthesized in G2 in low levels and reaches full loading in G1, generating a high fluorescent signal, usually in paired cells. Afterwards, ORC1b degrades rapidly at the at G1/S transition, being undetectable in S-phase (Vergara, 2017). In the SDG24.1-mRFP x ORC1b-GFP cross we observed ORC1b G1/S patterns colocalizing with SDG24.1 nucleoli along the RAM (**Fig. 4.17 A**).

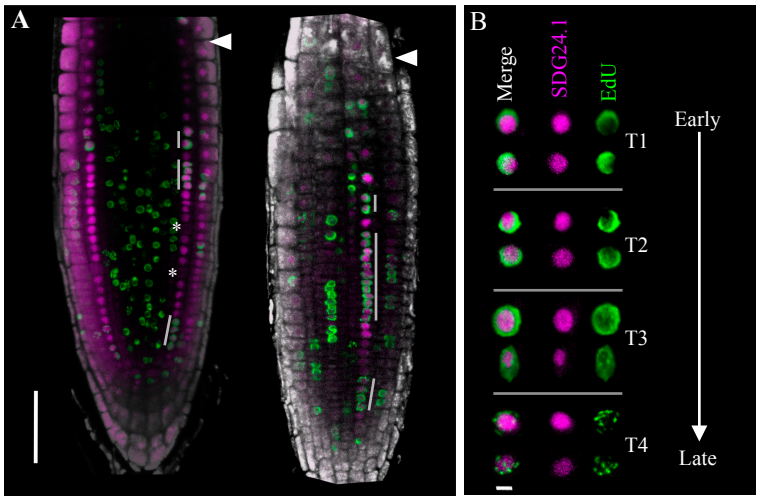


Figure 4.16. SDG24 is present in S-phase cells of the RAM. Roots expressing SDG24.1-mRFP (magenta) were labelled for 15 min with EdU (green) and detected by Click-it reaction. (**A**) Confocal images at the QC (left) or epidermis (right) planes showing SDG24.1 signal in S-phase cells (+EdU) along the RAM (grey lines). Arrowheads indicate the meristem end. Asterisks (*) point to negative SDG24.1/EdU cells. (**B**) Maximum intensity projections of nuclei from selected confocal *z*-stacks of different root cell layers showing individual EdU labelling types (T1-T4). DAPI was used as counterstaining (Gray). Bar= 50 µm in A, and 5 µm in B.

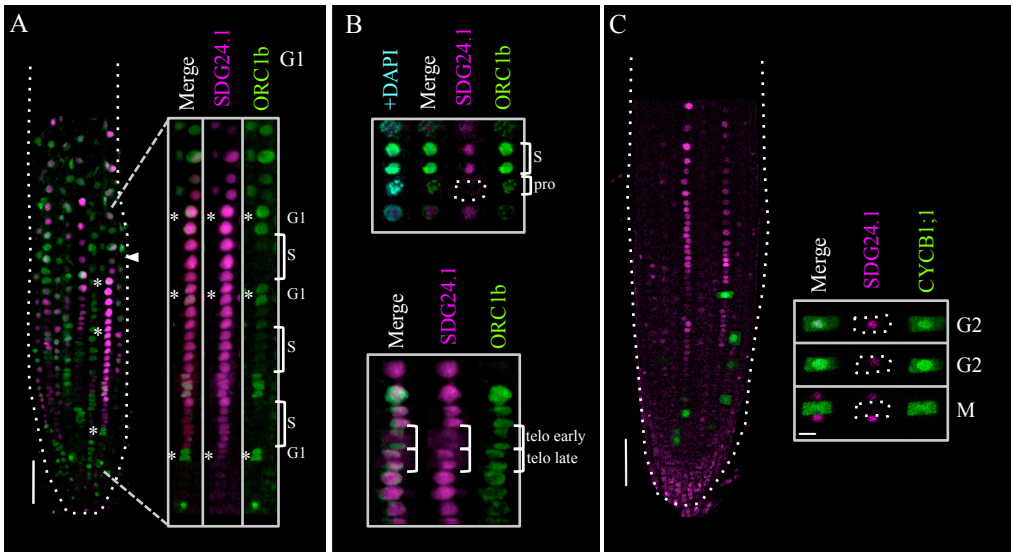


Figure 4.17. SDG24 protein is present in G1 and G2 phase cells in the RAM. Confocal images of 6 dps roots expressing SDG24.1-mRFP/ORC1b-GFP (**A, B**) or SDG24.1-mRFP/CYCB1;1-GFP in the nuclei (**C**). (**A**) Amplified epidermal cell file showing co-localization of SDG24.1 protein in cells at G1 (*) and S-phase (S, empty ORC1b expression) along the RAM. An arrowhead indicates the meristem end. Bar= 50 µm. (**B**) ORC1b marker also allows to confirm the absence of SDG24.1 at prophase nuclei (upper panel), and re-aggregation of the protein inside the nuclear area in different telophase stages (bottom panel) in the same cell file. All images are maximum intensity projections from selected confocal *z*-stacks. (**C**) Co-localization of SDG24.1-mRFP/CYCB1;1-GFP expression in cells at different G2/M stages confirm SDG24.1 degradation at the entering of mitosis. Bar= 50 µm (in A) and 10 µm (in C). pro: prophase; telo: telophase; M: mitosis; S: S-phase; G1/G2: Gap phases.

Other detailed observations allowed us to corroborate SDG24.1 absence in prophase nuclei and re-aggregation at telophases when ORC1b its still loading (**Fig. 4.17 B**). On the other hand, a cross with CYCB1;1-GFP, a cell marker of G2-M transition, allowed us to observe that the amount of SDG24.1 protein decreased as CYCB1;1 increases with progression from G2 into mitosis (**Fig. 4.17 C**).

Taken together these data, we confirmed our previous results of EdU labeling of S-phase cells and observations of mitotic figures. We conclude that during the cell cycle SDG24 expression starts to increase in telophase, is maintained in G1, S and until G2 phase, and is degraded at the entrance in prophase. Our results reinforce the use of crosses with ORC1b and CYCB1;1 markers as a useful way to study SDG24 cell cycle dynamics without introducing extra chemical stress.

4.1.3.3 SDG24 content along Arabidopsis root

After visual inspection of different roots expressing SDG24.1-mRFP, it became apparent that a portion of consecutive cells within the same cell file at the end of the RAM contained more fluorescent intensity than others. Therefore, to corroborate this observation we quantified the SDG24.1-mRFP fluorescent intensity of all nuclei in several epidermal cell files, from the first observed epidermal nucleus at the root tip to the first rapidly elongating cell at the TZ (Casamitjana-Martínez *et al.*, 2003). DAPI was used to contour the nuclear area and to estimate ploidy level. In the **Figure 4.18 A** is presented an example of one measured cell file.

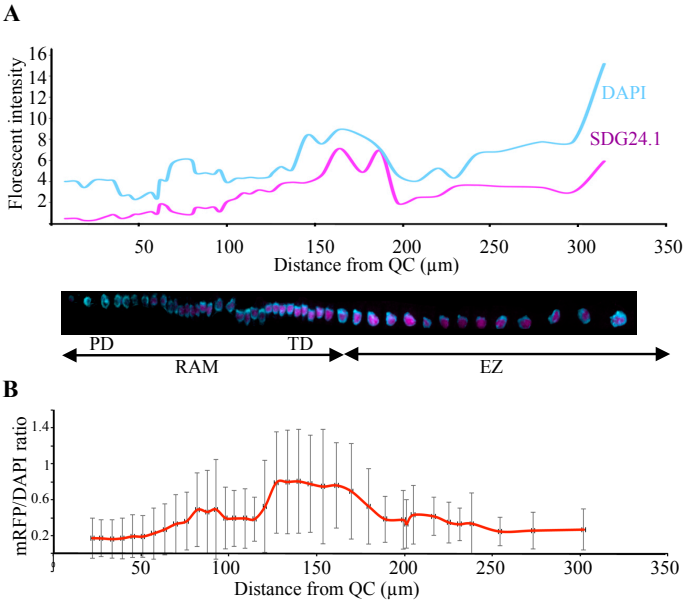


Figure 4.18. SDG24 content in the nuclei along the Arabidopsis root. (A) The relative fluorescence intensity of SDG24.1-mRFP (magenta) of each nucleus in entire cell files (n=10) was plotted against its position in the RAM (µm). Relative DNA was determined after DAPI (cyan) staining (ploidy level, Y-axis). An example of an epidermal cell file (trichoblast) is shown at the bottom, reconstructed from maximum intensity projections of confocal *z*-stacks. The RAM, including the proliferation domain (PD) and the transition domain (TD), as well as the elongation zone (EZ) are indicated. **(B)** Relative protein content as ploidy level increases along the root. Ratio (± SD, n=10) was measured by dividing relative fluorescent intensity of SDG24.1-mRFP by DAPI.

From this analysis, we observed a gradual increase of SDG24.1 fluorescence in nucleoli until it peaked at cells with 8C DNA content, at the end of meristem. Then, fluorescence diminished drastically in the TZ. This pattern is in agreement with initial observations made with he SDG24.1-GUS reporter line (**Fig. 4.7 B**). The ratio between protein content and ploidy level (measured by dividing fluorescent intensity of SDG24.1 by DAPI) clearly revealed this trend, where accumulation of SDG24.1 protein is preferentially high at the end of meristem and decreases at the entering of EZ (**Fig. 4.18 B**).

These results suggest that SDG24 expression, high at the end of the RAM and lower as differentiation takes place, might be due to a diminished nucleolar activity related to the lost of proliferative capacity of the cell in the EZ.

4.2 Functional analysis of mutants in *SDG7* and *SDG24* genes

4.2.1 Identification of *SDG7* and *SDG24* mutant alleles

To determine the biological function of SDG7 and SDG24 proteins, we obtained T-DNA insertion lines from the European Arabidopsis Stock Center (TAIR). Genomic PCR analyses were carried out to confirm insertions within the genomic region structure of each T-DNA line. During the course of this project, three *SDG7* mutant alleles were published, named: *sdg7-1*, *sdg7-2,* and *sdg7-3* (Lee *et al.*, 2015). In particular, mutant *sdg7-1* was described causing a missense mutation (Pro to Leu) in the SET domain, *sdg7-2* (Salk_131218) as a T-DNA knockout (KO) allele, and *sdg7-3* (Salk_072682) as a knockdown (KD)(**Fig. 4.19 A**). Additionally in our project three insertion lines were chosen, *sdg7-2* being one of them. Following the established nomenclature, the two other alleles that we identified were named *sdg7-4* (Salk_143603) and *sdg7-5* (WiscDLox430F09). All of them localized within introns, either at the beginning of gene (*sdg7-5*) or at the end of it (*sdg7-4*) (**Fig. 4.19 A**).

Transcript levels of *sdg7-2*, *sdg7-4 and sdg7-5* were evaluated by qPCR in three different regions before and after exons 5 and 7, a region that encode for the SET domain in the final protein structure (**Fig. 4.19 B**). Primers P5-6 and P10-11 were not able to amplify a band from *sdg7-5,* suggesting for a possible KO (or truncated protein), whereas a faint expression appeared for the *sdg7-4* allele, suggesting a possible KD. In the case of *sdg7-2,* amplification with P3-4 primer was also drastically diminished, in agreement with published data (Lee *et al.*, 2015).

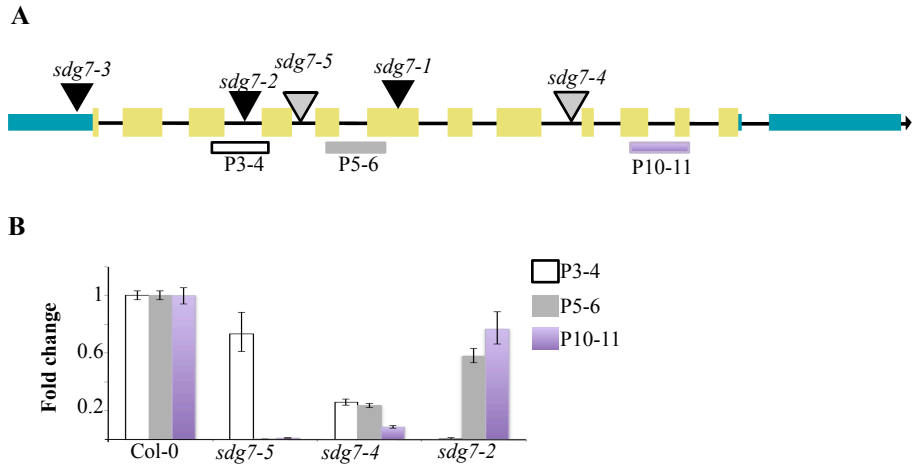


Figure 4.19. Relative expression level of *SDG7 SDG24* in *sdg7* mutant plants. (**A**) Disruption of *SDG7* gene by different T-DNA insertion lines. Black triangles indicate previously described mutant alleles: *sdg7-1*, *sdg7-2* (Salk_131218) and *sdg7-3* (Salk_072682). Gray triangles point to the insertion lines used in this project: *sdg7-4* (Salk_143603) and *sdg7-5* (Wisc430F09). Exons (yellow boxes), introns (black line) and 5′- 3′ UTR (green) are also indicated. **(B)** qPCR on cDNA from 7 dps seedlings. Histogram bars represent the fold change given by the $2^{-\Delta\Delta CT}$ method comparing wild type (Col-0) and *sdg7* mutants. . The position of primers used for qPCR is shown in **A**: P3-4 (white), P5-6 (grey) and P10-11 (purple). ± SD between technical replicas. At least 2 biological replicas were performed with similar results. *ACT2* served as gene of reference.

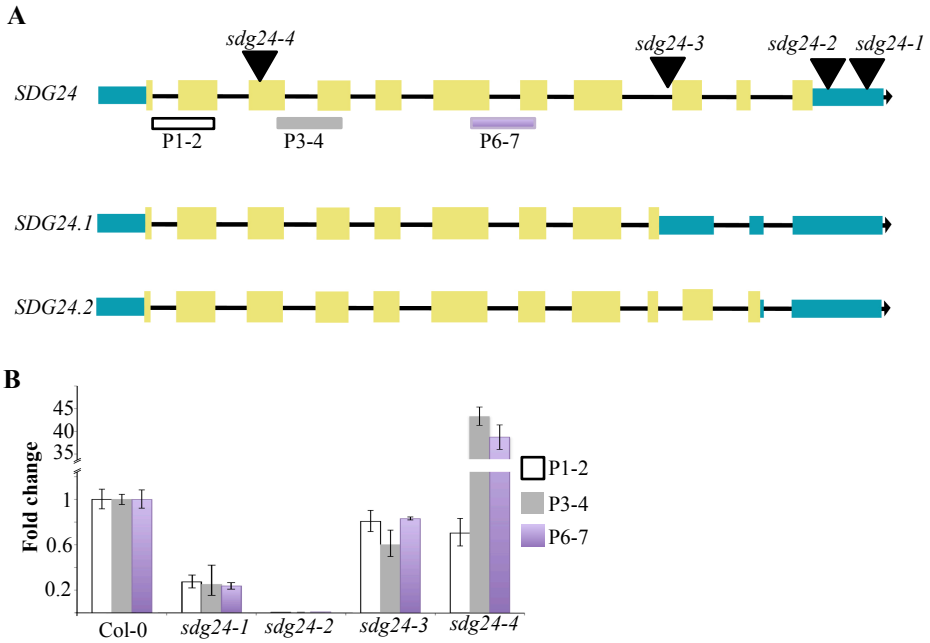


Figure 4.20. Relative expression level of *SDG24* in *sdg24* mutant plants. (A) Disruption of *SDG24* gene by T-DNA insertion lines. Alternatively spliced variants are indicated below (*SDG24.1* and *SDG24.2*). Triangles indicate the insertion site of *sdg24-1* (Wisc492G18), *sdg24-2* (Salk_201808), *sdg24-3* (Salk_205054) and *sdg24-4* (SK22803). Exons (yellow boxes), introns (black line) and 5′- and 3′-UTR (green musk) are displayed. (**B**) qPCR on cDNA from flower buds. Histogram bars represent the fold change given by the $2^{-\Delta\Delta CT}$ method comparing the wild type (Col-0) and *sdg24* mutants. The position of primers used for qPCR is shown in **A**: P1-2 (white), P3-4 (grey) and P6-7 (purple). ± SD between technical replicas. At least 2 biological replicas were performed with similar results. *ACT2* served as gene of reference.

In the case of *SDG24*, several T-DNA lines were also available. Here, we evaluated four candidates*:* *sdg24-1* (WiscDsLOX489-492G18), *sdg24-2* (Salk_201808), *sdg24-3* (Salk_205054) *and sdg24-4* (SK22803). In the **figure 4.20 A** the location of different alleles in the *SDG24* gene is described: one inside exon III (*sdg24-4*) and three at the 3′ region (*sdg24-1/2/3*). qPCR analysis revealed that untranslated 3′ exons are important to produce stable transcript levels, as *sdg24-1* (KD) and *sdg24-2* (KO) alleles showed a drastic reduction of expression in all regions tested; however, this effect was not observed for *sdg24-3*. On the contrary, the insertion caused by *sdg24-4* increased *SDG24* expression at high levels, possibly acting as a gain-of-function allele (**Fig. 4. 20 B**).

With the results gathered by the transcriptional analysis, we considered the use of *sdg7-4*, *sdg7-5*, *sdg24-1* and *sdg24-2* mutants alleles for the consecutive functional analysis.

4.2.2 Enzymatic activity of SDG7 and SDG24

4.2.2.1 The H4K20 methylation mark in *sdg7* and *sdg24* mutants

The homology with known lysine HMTases and the nuclear localization of SDG7 and SDG24.1 led us to investigate the possible changes of H4K20 methylation pattern in *sdg7* and *sdg24* mutants. To do so, we performed Western blot analysis on histone-enriched protein extracts from seedlings of *sdg7-4, sdg7-5, sdg24-1* and *sdg24-2*, and compared them with the wild type (Col-0) (**Fig. 4.21 A**). Unfortunately, the H4K20me3 antibody described in the literature (Xu *et al.*, 2008; Sanchez and Gutierrez, 2009) failed to detect this mark in our Western blot experiments, and moreover, showed many unspecific bands.

We also tested other conditions, including increasing the amount of nuclear protein, different histone protocol extractions, other specific tissues (roots), and several H4K20me3 antibodies (detailed in Materials and Methods), but results were inconclusive. Likewise, detection of H4K20me1 used as H4 reference also showed high unspecificity. As a way to overcome the failure in detecting H4K20me3, we tested the possibility that H4K20 was modified at different time points of the cell cycle, a situation described for H4K20me1 in humans (Jorgensen *et al.*, 2013). Thus, we synchronized Arabidopsis MM2d cultured cells by sucrose starvation and released them at G0-G1 (Menges and Murray, 2002). Samples were taken at different time points of the cell cycle and subjected to Western blot. Once again, we did not detect the H4K20me3 mark (**Fig. 4.21 B**).

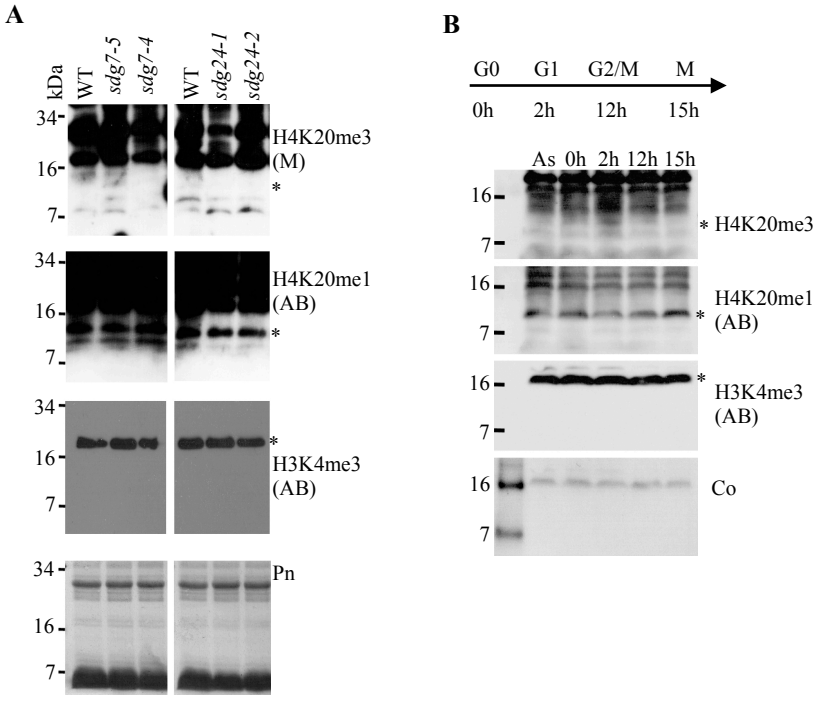


Figure 4.21. H4K20me3 methylation pattern in *sdg7* and *sdg24* mutants. (**A**) H4K20me3 methylation pattern in wild type (Col-0), *sdg7* and *sdg24* mutant lines, using H4K20me1 and H3K4me3 as controls. 20 µg of nuclear protein extract from 10 dps seedlings were subjected to Western blot (Bonenfant *et al.*) analysis using antibodies from Abcam (AB) or Merck-Millipore (M), as indicated. (**B**) H4K20me3 pattern during cell cycle after sucrose starvation-induced synchrony of MM2d cells. Exponential-phase Arabidopsis MM2d cells were subjected to sucrose removal and resupply to follow cell cycle from G0 to mitosis. Nuclear protein extracts from asynchronous cells (As) and synchronized cells at different time points after release (0 to 15 hours) were analysed by WB. 250 µg of protein were loaded. A diagram above depicts the approximate cell cycle phases at different times after synchronization. Molecular size markers (kDa) are displayed at the left of each image. Coomassie (Co) or Ponceau (Pn) stained membranes are shown at the bottom. The H3 and H4 location is indicated by asterisks (*).

From these data, we conclude that: (i) the lack of specificity of current available antibodies did not allow us to identify the H4K20me3 mark by Western blot neither in the wild type nor in *sdg7* and *sdg24* mutant lines; and (ii) SDG7 and SDG24 might have specificity for other histone modification.

4.2.2.2 Evaluation of alternative methylation marks for SDG7 and SDG24

In order to test alternative histone residues being modified by the action of SDG7 and SDG24, we performed a peptide binding assay and an *in vitro* histone methyltransferase assay. The first technique allows the multiple screening of posttranslational modifications in histone tails that are recognized by the proteins of interest, and the second approach allows the detection of modifications using *S*-(5′-Adenosyl)-*L*-methionine (SAM). To do so, SDG7 and SDG24.1 coding sequences were fused to the glutathione S-transferase (GST) or 6-histidine residues (His6X) for bacterial purification (**Fig. 4.22 A**).

When expressed in bacteria under standard conditions (0.6 OD, induction with 0.5 mM IPTG for 1 h at 30 ºC) both GST-SDG7 and His6X-SDG7 tagged proteins co-purified with the insoluble fraction; on the contrary, SDG24.1 tagged protein was soluble. In order to obtain more soluble protein, we optimized the induction and purification conditions of His6X-SDG7, and the same conditions were applied for His6X-SDG24.1 (**Fig. 4.22 B, C**), which allowed the purification of sufficient protein in the soluble fraction. Modifications included the addition of glucose to the starter culture (0.25%), lower amount of bacteria to be induced (0.4 OD), a reduced IPTG concentration (0.3 mM) and a prolonged induction time at a lower temperature (2.5 hours at 18 ºC). Purification steps proceeded in an affinity Ni-NTA matrix, with some additional modifications over the manufacturer's standard conditions, including an increased NaCl concentration in the Lysis buffer (800 mM), lower imidazole concentration (5 mM) and the addition of non-ionic detergent (1% triton).

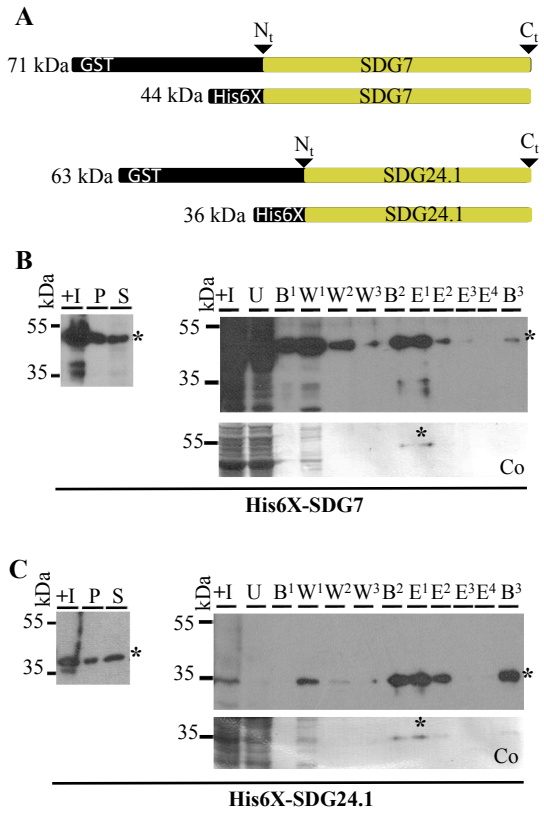


Figure 4.22. Expression of recombinant SDG7 and SDG24 proteins in bacteria. (**A**) *SDG7* and *SDG24.1* coding sequences were fused to GST or His6X tags and expressed in *E. coli* (SDG7, 41 kDa, and SDG24.1, 33 kDa). The expected molecular mass of the fused proteins is shown at the left of each protein diagram. (**B**) Optimized induction and purification conditions of soluble His6X-SDG7. Purification was under native conditions with Ni-NTA agarose matrix. Each purification step was analysed by WB (right panel). Lower panel shows Coomassie stained membrane (Co) used for WB. (**C**) Same conditions as in **B**, but for His6X-SDG24.1 protein. Asterisks (*) indicate the protein location. +I, sonicated IPTG induced culture; S, soluble fraction; P, insoluble fraction; U, unbound proteins to Ni-NTA matrix; B^1, proteins bound to matrix; W^1-W^3, washing steps; B^2, proteins bound to matrix after the washing steps; E^1-E^4, imidazole elution steps; B^3, proteins bound to matrix after the elution steps. Molecular markers (kDa) are shown to the left of each image. α-His and α-GST antibodies where used to detect the corresponding protein tags.

4.2.2.2.1 SDG7 and SDG24 bind to histones *in vitro*

Although our candidate proteins were selected for their homology to functional H4K20me3 HMTases, e.g. the mammalian SUV4-20H1/H2 enzymes, mutations in SDG7 and SDG24 did not alter H4K20me3 methylation levels, leading to hypothesize a diverse substrate

specificity for both proteins. Thus, we performed a histone peptide-binding assay (MODified Histone Peptide Array, Active Motif) where SDG7 and SDG24.1 were tested for their capacity to recognize different histone marks (**Fig. 4.23**).

Peptide arrays contained 59 different post-translational modifications for histone acetylation, methylation, phosphorylation, and citrullination on the N-terminal tails of histones H2A, H2B, H3 and H4, with 384 combinations, by duplicate. For the assay, His6X-SDG7 and His6X-SDG24.1 proteins were expressed in *E. coli* (**Fig. 4.23 A**) and incubated with the peptide arrays. The binding signal was detected by Western blot against the His6X tag (**Fig. 4.23 B**), and analyzed with the aid of the array analysis software (Active Motif). Measurements were given by the activity of spot intensity, calculated by averaging individual spotted peptides of both sides of the slide. To identify the strongest interactions, we elaborated a table in which individual and multiple PTMs are distributed in the four histones (**Tables 4.1 and 4.2**). In the Appendix 1 are shown the average of activity of spot intensity obtained for all 384 combinations with His6X-SDG7 and His6X-SDG24.1.

In the case of His6X-SDG7, assay showed binding capacity to lysine (K) and arginine (R) residues of histone H3, including acetylated and methylated marks such as H3K4ac, H3K9me3, H3K27me1, H3R26me2 and H3R17me2 (**Table 4.1**). A combination of two, three or four marks including different arrangements of H3R2me2, H3R8me2, H3K4me1/me2/me3, H3K4ac, H3K9me1/me2/me3 and H3K9ac also gave high intensity scores. SDG7 also showed binding capacity to histone H4 peptides carrying one modification, corresponding to R3me2, S1P, K5ac, K16ac. Moreover, binding for histones H2A and H2B with highest intensity scores, bearing K5ac, S1P and K9ac for the case of H2A; and, S14P, K5ac, K12ac and K15ac marks for histone H2B.

The binding pattern of His6X-SDG24.1 to histones H3, H4, H2A and H2B (**Table 4.2**) was quite different. For histone H3, motifs including H3K27me2/me3 and H3R2me2 were more represented, and signals with a combination of histone H3 motifs R2me2, K4me2/me3, R8me2 and K9ac was also detected. Interestingly for histone H4, peptides including K20me1/me2/me3 marks presented a high detection signal. In the case of H2A, a combination of acetylated of K13ac, K9ac, K5ac and K13ac seems to have a mayor contribution for peptide binding. For histone H2B also an acetylated pattern was detected, including peptide marks combining K5ac, K12ac and K15ac residues.

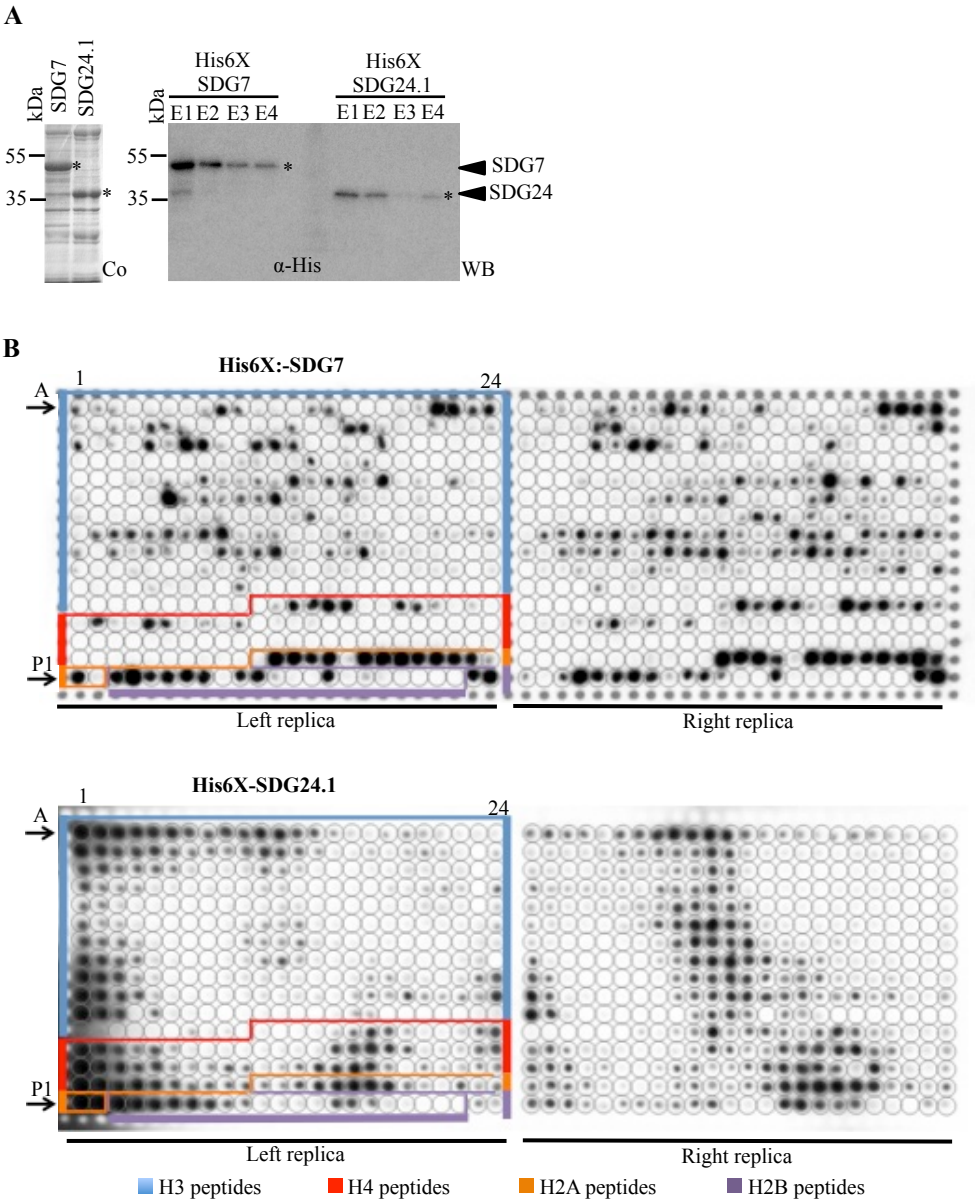


Figure 4.23. Histone peptide binding assay of SDG7 and SDG24 proteins. (**A**) Purification of recombinant His6X-SDG7 and His6X-SDG24.1 proteins used in this assay, from 1 L of bacterial culture. Left panel shows Coomassie blue-stained SDS-PAGE gel (Co) of purified proteins (*). Right panel displays WB analysis (α-His antibody) of four elution steps (E1-E4) of each protein after dialysis. For the histone peptide-binding assay the fractions E1-E4 were pulled together. Molecular markers (kDa) are shown to the left. (**B**) MODified Histone Peptide Arrays (CelluSpots, Active Motif) were incubated with either His6X-SDG7 or His6X-SDG24.1 recombinant proteins overnight. Slides were analysed by Western blot with α-His antibody. Wells (A1-P24) are countered to indicate the histone peptides combinations corresponding to histones H3 (blue), H4 (red), H2A (orange) and H2B (purple).

Table 4.1. His6X-SDG7 Histone peptide binding results

No.	Loc	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
9	A9	H3 1-19	K4ac				0.849
14	A14	H3 1-19	K9me2				0.579
240	J24	H3 16-35	K27me1				0.488
238	J22	H3 16-35	R26me2a				0.466
224	J8	H3 7-26	R17me2s				0.436
11	A11	H3 1-19	R8me2a				0.422
10	A10	H3 1-19	R8me2s				0.331
237	J21	H3 16-35	R26me2s				0.328
21	A21	H3 1-19	R2me2s	K4me1			0.997
55	C7	H3 1-19	R8me2a	K9me3			0.977
22	A22	H3 1-19	R2me2s	K4me2			0.957
56	C8	H3 1-19	R8me2a	K9ac			0.929
24	A24	H3 1-19	R2me2s	K4ac			0.896
59	C11	H3 1-19	R8Citr	K9me1			0.817
23	A23	H3 1-19	R2me2s	K4me3			0.761
66	C18	H3 1-19	K9me1	T11P			0.708
53	C5	H3 1-19	R8me2a	K9me1			0.645
29	B5	H3 1-19	R2me2a	K4ac			0.640
48	B24	H3 1-19	R8me2s	K9me2			0.596
30	B6	H3 1-19	R2Citr	T3P			0.573
60	C12	H3 1-19	R8Citr	K9me2			0.533
54	C6	H3 1-19	R8me2a	K9me2			0.420
250	K10	H3 16-35	R26me2a	K27me1			0.370
61	C13	H3 1-19	R8Citr	K9me3			0.345
114	E18	H3 1-19	R2me2a	K4me3	K9me2		0.993
132	F12	H3 1-19	K4me1	R8me2s	K9me2		0.889
109	E13	H3 1-19	R2me2s	K4me2	K9me1		0.828
112	E16	H3 1-19	R2me2a	K4me1	K9me2		0.643
118	E22	H3 1-19	R2me2s	K4me3	K9me3		0.637
130	F10	H3 1-19	K4me3	R8me2a	K9me1		0.600
116	E20	H3 1-19	R2me2s	K4me1	K9me3		0.565
126	F6	H3 1-19	K4me3	R8me2s	K9me1		0.524
41	B17	H3 1-19	R2me2s	T3P	K4me3		0.500
129	F9	H3 1-19	K4me2	R8me2a	K9me1		0.500
138	F18	H3 1-19	K4me3	R8me2a	K9me2		0.491
131	F11	H3 1-19	K4ac	R8me2a	K9me1		0.477
128	F8	H3 1-19	K4me1	R8me2a	K9me1		0.436
111	E15	H3 1-19	R2me2s	K4ac	K9me1		0.435
86	D14	H3 1-19	R8me2s	K9me3	T11P		0.420
110	E14	H3 1-19	R2me2s	K4me3	K9me1		0.360
104	E8	H3 1-19	R2me2a	K4me1	R8me2a		0.354
204	I12	H3 1-19	R2me2s	K4me1	R8me2a	K9me3	0.953
177	H9	H3 1-19	R2me2a	K4me2	R8me2s	K9me3	0.946
209	I17	H3 1-19	R2me2a	K4me2	R8me2a	K9me3	0.923
176	H8	H3 1-19	R2me2a	K4me1	R8me2s	K9me3	0.916
201	I9	H3 1-19	R2me2a	K4me2	R8me2a	K9me2	0.859
174	H6	H3 1-19	R2me2s	K4me3	R8me2s	K9me3	0.854
203	I11	H3 1-19	R2me2a	K4ac	R8me2a	K9me2	0.776
172	H4	H3 1-19	R2me2s	K4me1	R8me2s	K9me3	0.775
184	H16	H3 1-19	R2me2a	K4me1	R8me2s	K9ac	0.770
202	I10	H3 1-19	R2me2a	K4me3	R8me2a	K9me2	0.728
158	G14	H3 1-19	R2me2s	K4me3	R8me2s	K9me1	0.718
171	H3	H3 1-19	R2me2a	K4ac	R8me2s	K9me2	0.693
191	H23	H3 1-19	R2me2s	K4ac	R8me2a	K9me1	0.617
207	I15	H3 1-19	R2me2s	K4ac	R8me2a	K9me3	0.606
197	I5	H3 1-19	R2me2s	K4me2	R8me2a	K9me2	0.584
180	H12	H3 1-19	R2me2s	K4me1	R8me2s	K9ac	0.577
192	H24	H3 1-19	R2me2a	K4me1	R8me2a	K9me1	0.558

Table 4.1. ***Cont.***

H4 peptides

Pep.	Loc	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
279	L15	H4 1-19	R3me2a				0.987
277	L13	H4 1-19	S1P				0.972
280	L16	H4 1-19	K5ac				0.942
283	L19	H4 1-19	K16ac				0.934
278	L14	H4 1-19	R3me2s				0.845
276	L12	H4 1-19	unmod				0.270
284	L20	H4 1-19	S1P	R3me2s			0.774
294	M6	H4 1-19	K12ac	K16ac			0.745
293	M5	H4 1-19	K8ac	K16ac			0.742
285	L21	H4 1-19	S1P	R3me2a			0.616
287	L23	H4 1-19	R3me2s	K5ac			0.569
290	M2	H4 1-19	R3me2a	K8ac			0.426
286	L22	H4 1-19	S1P	K5ac			0.372
288	L24	H4 1-19	R3me2s	K8ac			0.262
298	M10	H4 1-19	R3me2a	K5ac	K8ac		0.581
296	M8	H4 1-19	S1P	R3me2a	K5ac		0.418

H2A peptides

Pep.	Loc	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
350	O14	H2a 1-19	K5ac				0.992
349	O13	H2a 1-19	S1P				0.990
351	O15	H2a 1-19	K9ac				0.942
348	O12	H2a 1-19	unmod				1.000
354	O18	H2a 1-19	S1P	K9ac			1.000
357	O21	H2a 1-19	K5ac	K13ac			0.999
353	O17	H2a 1-19	S1P	K5ac			0.999
355	O19	H2a 1-19	S1P	K13ac			0.995
358	O22	H2a 1-19	K9ac	K13ac			0.992
356	O20	H2a 1-19	K5ac	K9ac			0.991
359	O23	H2a 1-19	S1P	K5ac	K9ac		0.975
361	P1	H2a 1-19	S1P	K9ac	K13ac		0.634
360	O24	H2a 1-19	S1P	K5ac	K13ac		0.606
363	P3	H2a 1-19	S1P	K5ac	K9ac	K13ac	0.894

H2B peptides

Pep.	Loc	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
367	P7	H2b 1-19	S14P				0.994
366	P6	H2b 1-19	K12ac				0.981
365	P5	H2b 1-19	K5ac				0.944
368	P8	H2b 1-19	K15ac				0.939
364	P4	H2b 1-19	unmod				1.000
371	P11	H2b 1-19	K5ac	K15ac			0.899
370	P10	H2b 1-19	K5ac	S14P			0.799
375	P15	H2b 1-19	K5ac	K12ac	S14P		0.925

Pep. No.= Peptide number in one replica grid (left or right). Location= position of peptide on the grid. Mod= modified mark in the peptide. Norm. act.= normalized activity of spot intensity.

Table 4.2. His6X-SDG24.1 histone peptide binding results

H3 peptides

Pep. No.	Loc	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
241	K1	H3 16-35	K27me2				0.862
2	A2	H3 1-19	R2me2s				0.744
242	K2	H3 16-35	K27me3				0.573
11	A11	H3 1-19	R8me2a				0.451
12	A12	H3 1-19	R8Citr				0.368
9	A9	H3 1-19	K4ac				0.299
7	A7	H3 1-19	K4me2				0.229
35	B11	H3 1-19	T3P	K4me1			0.670
235	J19	H3 7-26	K14ac	R17me2a	K18ac		0.263
217	J1	H3 1-19	R2me2a	K4me2	R8me2a	K9ac	0.822
193	I1	H3 1-19	R2me2a	K4me2	R8me2a	K9me1	0.677
218	J2	H3 1-19	R2me2a	K4me3	R8me2a	K9ac	0.622
169	H1	H3 1-19	R2me2a	K4me2	R8me2s	K9me2	0.402

H4 peptides

Pep. No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
313	N1	H4 11-30	K20me1				0.968
314	N2	H4 11-30	K20me2				0.711
315	N3	H4 11-30	K20me3				0.465
281	L17	H4 1-19	K8ac				0.444
306	M18	H4 11-30	unmod				0.435
307	M19	H4 11-30	K12ac				0.238
308	M20	H4 11-30	K16ac				0.233
337	O1	H4 11-30	R19me2s	K20me2			0.951
338	O2	H4 11-30	R19me2s	K20me3			0.839
339	O3	H4 11-30	R19me2s	K20ac			0.608
290	M2	H4 1-19	R3me2a	K8ac			0.594
326	N14	H4 11-30	K16ac	K20me3			0.373
327	N15	H4 11-30	K16ac	K20ac			0.360
325	N13	H4 11-30	K16ac	K20me2			0.203
328	N16	H4 11-30	K12ac	K16ac	K20me1		0.625
329	N17	H4 11-30	K12ac	K16ac	K20me2		0.480
305	M17	H4 1-19	K5ac	K8ac	K12ac	K16ac	0.521
304	M16	H4 1-19	R3me2a	K5ac	K8ac	K12ac	0.346
303	M15	H4 1-19	R3me2s	K5ac	K8ac	K12ac	0.329

H2A peptides

Pep. No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
352	O16	H2a 1-19	K13ac				0.808
351	O15	H2a 1-19	K9ac				0.522
350	O14	H2a 1-19	K5ac				0.440
353	O17	H2a 1-19	S1P	K5ac			0.711
354	O18	H2a 1-19	S1P	K9ac			0.606
361	P1	H2a 1-19	S1P	K9ac	K13ac		1.000
362	P2	H2a 1-19	K5ac	K9ac	K13ac		0.972
363	P3	H2a 1-19	S1P	K5ac	K9ac	K13ac	0.868

H2B peptides

Pep. No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
375	P15	H2b 1-19	K5ac	K12ac	S14P		0.817
376	P16	H2b 1-19	K5ac	K12ac	K15ac		0.812

Pep. No.= Peptide number in one replica grid (left or right). Location= position of peptide on the grid. Mod= modified mark in the peptide. Norm. act.= normalized activity of spot intensity.

From the outcome of the combinatorial peptide-binding assay we conclude that both SDG7 and SDG24 posses –at least– an *in vitro* capacity to bind to different histone PTMs, being those in a single or a grouped manner. From these results we can hypothesize that: (i) SDG7 and SDG24 might be able to identify different chromatin scenarios, and (ii) they possibly have different histone substrate specificity.

4.2.2.2.2 Histone methyltransferase activity of SDG7 and SDG24 *in vitro*

Based on the peptide-binding assay we wanted to test the possible methylation activity of SDG7 and SDG24.1 on some of the more likely residues of H3 and H4 directly by performing HMTase assays. Recombinant His6X-SDG7 and His6X-SDG24.1 were expressed in *E. coli* and incubated with core histones from calf thymus and SAM (**Fig. 4.24 A-C**). Unfortunately, this assay did not identify any relevant activity HMTase activity with the antibodies tested (**Fig. 4.24 D**).

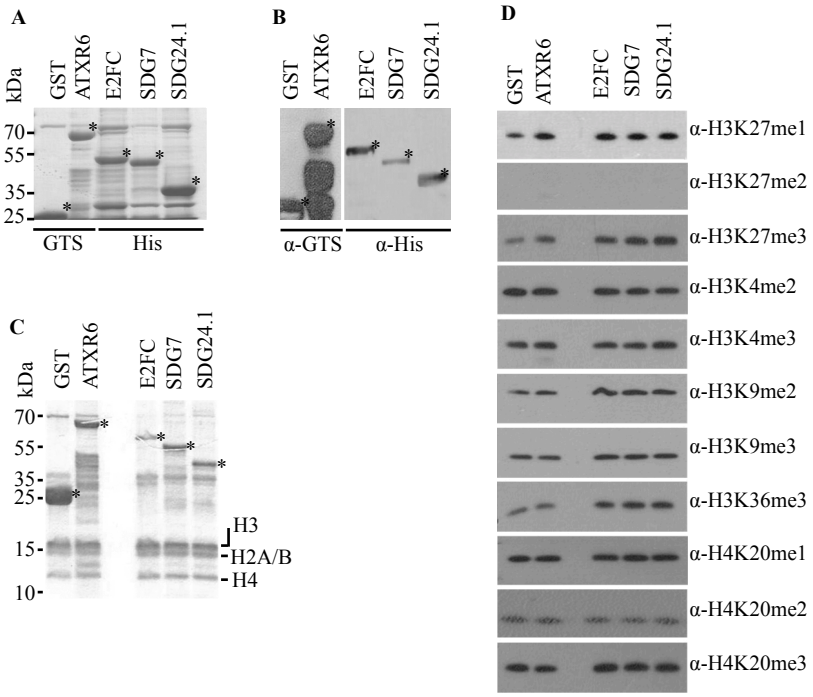


Figure 4.24 Histone methyltransferase assay of SDG7 and SDG24 proteins. (**A**) Coomassie blue-stained SDS-PAGE of purified recombinant proteins His6X-SDG7, His6X-SDG24.1, GST-ATXR6$_{PHD-SET}$, GST and His6X-E2FC used in this assy. GST from pGEX-6P and His6X-E2FC were used as background controls for GST and His6X tagged proteins, respectively. GST-ATXR6$_{PHD-SET}$ (Jacob et al., 2009) was employed as a positive control for HMTase enzymatic activity. (**B**) WB analysis using antibodies α-GTS and α-His to detect the purified proteins shown in panel A. (**C**) Coomassie blue-stained SDS-PAGE gel of *in vitro* methylation assay with purified recombinant proteins and calf thymus histones. (**D**) Western blot analysis of the methylation assay shown in panel C. Molecular markers (kDa) and protein location (*) are shown.

Negative results could be due to unfavorable reaction conditions. Although we performed alternative incubation settings, including changes in pH, temperature, incubation time, and antibodies from different commercial sources; all of them showed inconclusive results (see Materials and Methods details). As a control, the reaction activity of GST-ATXR6 recombinant protein showed positive methylation of H3K27me1, as expected, although without displaying the drastic intensity variation previously described (Jacob *et al.*, 2009). Also, in general we observed that histone background was high for all the antibodies tested (except for H3K27me2), suggesting that the histone already was significantly modified.

From the results of HMTase assay we were not able to confirm a methyltransferase activity of SDG7 or SDG24.1 towards histone H3 or H4. Nevertheless, we do not discard their possible activity as other approaches can be performed to identify alterations of histone methylation states, including the use of plant recombinant histones, reconstituted nucleosomes, and also a histone PTMs screening by Mass spectrometry in mutant alleles.

4.2.3 Phenotypical characterization of *sdg7* and *sdg24* mutants during development

Plant organs have to grow to a certain size, and different regulatory mechanisms involving intrinsic and environmental cues are required to enable optimal organ development (Sparks *et al.*, 2013; Czesnick and Lenhard, 2015). Despite our efforts to associate HMTase activity *in vitro* for both SDG7 and SDG24.1 variant, we evaluated possible growth alterations in the mutant lines. As described previously in this work, SDG7 and SDG24 are expressed at different developmental stages.

4.2.3.1 Aerial plant development is altered in *sdg7* and *sdg24* mutants

Arabidopsis is a facultative long day (LD) plant whose flowering is delayed under short day (SD) conditions. When SD occurs, vegetative program continued for a longer period, thereby changing plant morphology (Jeong and Clark, 2005; Rosas *et al.*, 2014; Cho *et al.*, 2017). Here, we tested if the *sdg7* and *sdg24* mutations could interfere with the aerial architecture controlled by photoperiod.

Homozygous plants for *sdg7-5* and *sdg24-2* single mutants, and *sdg7-5, sdg24-2* double mutant were grown under LD and SD conditions (**Fig. 4.25 A**, **B**). Both *sdg7-5* and *sdg24-2* were smaller than the wild type plants (Col-0) especially when grown under SD. This phenotype seems to be related with organ size, as the flowering time did not showed differences under LD nor SD (**Fig. 4.25 E**, **F**).

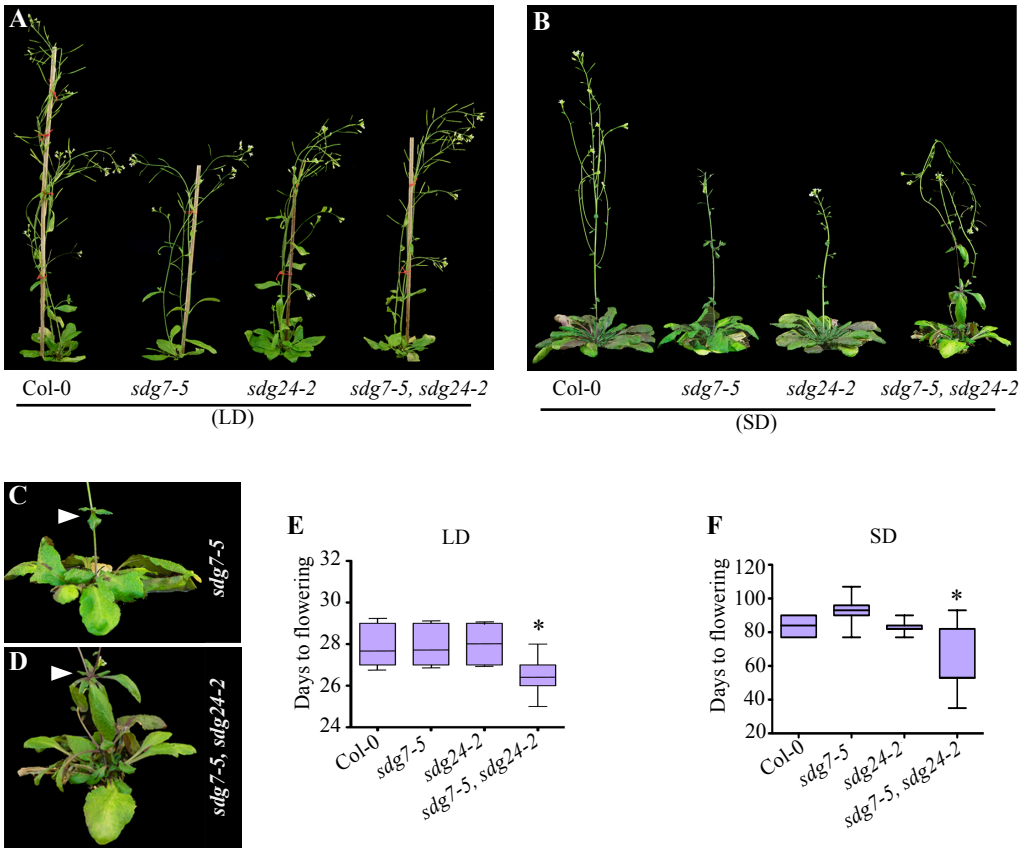


Figure 4.25. Growth of *sdg7* and *sdg24* mutant plants. (A) Wild type (Col-0), *sdg7-5, sdg24-2 and sdg7-5, sdg24-2* double mutant plants grown under long day (LD) or short day (SD) growth conditions, as indicated. Images were taken at 45 dps for LD, and 90 dps in SD. **(C, D)** A secondary rosette phenotype is indicated by white arrowheads in the *sdg7-5 and sdg7-5, sdg24-2* double mutant. **(E-F)** Flowering time quantified as the number of days to produce flower buds under LD **(E)** or SD **(F)** conditions. Error bars indicate the SD. Statistical significance between genotypes, (*) $P<0.01$, was calculated by the ANOVA-Dunnett's multiple comparisons test, n=15 plants.

On the contrary, in the case of *sdg7-5, sdg24-2* double mutant a significant early flowering time phenotype was observed in both photoperiods, with a greater difference detected at SD, giving to the plant a taller appearance than parental lines. This also suggests some genetic interactions between the two mutations.

We also measured the ploidy levels in rosette leaves (L1-L2 and L3-L4 pairs) and cotyledons at 30 dps under LD conditions, but no alteration was observed between the wild type and mutant lines (data not shown). Intriguingly, a particular phenotype emerged under SD in both *sdg7-5* and *sdg7-5, sdg24-2* double mutant, such as the formation of secondary rosette-like structures, namely aerial rosettes, detectable between the 1st and 2nd stem internode (**Fig. 4.25 C**, **D**). This phenotype likely associated with the *sdg7-5* mutation.

After plants reached the end of their life cycle, despite plant size phenotype, all mutant lines were able to produce high amount of seeds, suggesting that fertility was not compromised.

4.2.3.2 Root growth is affected in *sdg7* and *sdg24* mutants

Root architecture has implications above the overall plant architecture, nutrient uptake and environmental plasticity, which will procure plant survival after different stresses (Petricka *et al.*, 2012; Jung and McCouch, 2013). In this work we found that SDG7 and SDG24 were expressed in the primary root and during LRP formation, suggesting a possible altered root morphology of their mutants. To corroborate this hypothesis, we measured primary root growth of mutant lines under LD and SD conditions for several days. Under both photoperiods, the primary root of all mutants was significantly smaller than wild type (**Fig. 4.26 A**, **B**, **E**). A 12-days time-course experiment corroborated alterations in the total root elongation of the mutants and wild type, where differences started to visualize at 6 dps under LD (**Fig. 4.26 C**), and between 9-10 dps for SD photoperiod (**Fig. 4.26 D**).

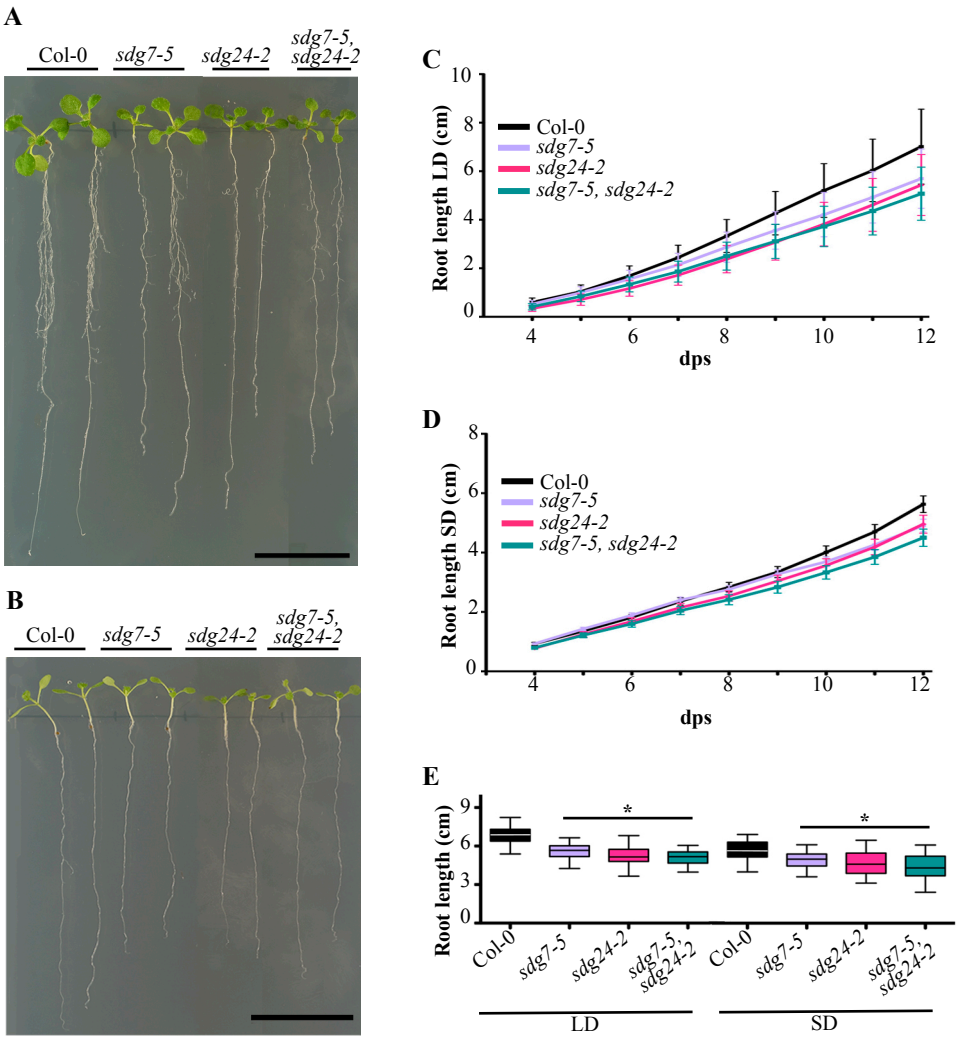


Figure 4.26 Principal root growth in *sdg7-5* and *sdg24-2* mutants. (A, B) Phenotype of seedlings of wild type (Col-0) and mutant lines under LD (**A**) or SD (**B**) growth conditions. Scale bars = 2 cm. Images were taken after 12 dps. Root growth was measured under LD (**C, E**) or SD conditions (**D, E**). Seeds were plated on Murashige and Skoog (MS) agar plates and counted for germination at the 3rd day post sowing (dps). Measurements were taken between 4-12 dps in cm. **(E)** Total root length of **C** and **D** at 12 dps. ± SD. Statistical significance (*) between mutants and WT, $p<0.01$, was calculated by the ANOVA-Dunnett's multiple comparisons test, n=21 plants for each genotype. Two biological replicas were made with similar results.

Additionally, during observation in LD conditions we perceived a reduction of lateral roots (LR) emergence in the *sdg7-5* and *sdg24-2* mutants (**Fig. 4.26 A**). We analyzed this by measuring LR density and LRP distribution. LR density was calculated by dividing the number of LR by the primary root length for each seedling. Mutants showed a reduction of LR at 10 dps, but only LR density of *sdg24-2* was significant among them (**Fig. 4.27 A**). To determine when the unbalanced LR growth occurred we analyzed the developmental stage of LRPs at 6 dps. At this time point, *sdg7-5, sdg24-2* single mutants and *sdg7-5, sdg24-2* double mutant accumulated around 10-15% more of early stages (StI- StIV) than wild type (**Fig. 4.27 B**), suggesting that the development has a slower progression in the mutant lines.

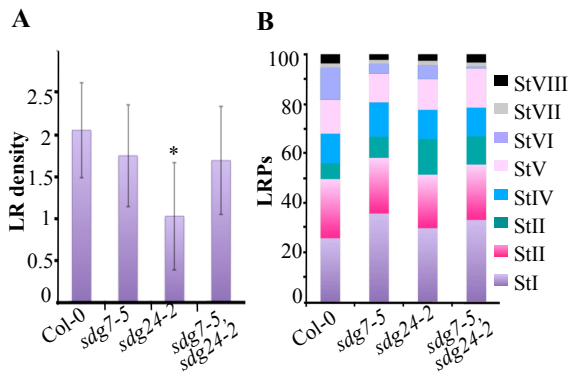


Figure 4.27. Lateral root development in *sdg7-5* and *sdg24-2* mutants. (**A**) Emerged lateral root density in wild type (Col-0) and mutant lines. Measurements were taken 10 dps on MS agar plates under LD conditions. Lateral root density was calculated by dividing the number of lateral roots by the primary root length for each seedling (20-27 seedlings per genotype). The density is reported as the number of lateral roots per cm of primary root. ± SD. (*) $P<0.01$, calculated by the ANOVA-Dunnett's multiple comparisons test. (**B**) Distribution of LRP at different stages in 6 dps roots grown under LD. LRP states (St) from I to VIII are shown. n=10 roots and ~100 LRPs scored.

4.2.3.3 The root meristem is affected in *sdg7* and *sdg24* mutants

After detection of an altered root growth in *sdg7* and *sdg24* mutants we wondered if an abnormal RAM development could be causing this phenotype. To do so, first we quantified the number and length of cortical cells, from QC to EZ, considering the end of the meristem the first rapidly elongated cell at the TZ (Casamitjana-Martínez *et al.*, 2003). Meristem analysis confirmed our hypothesis, showing smaller meristem size in the *sdg7-5, sdg24-2* single mutants and in *sdg7-5, sdg24-2* double mutant at 6 dps (**Fig. 4.28 A**, **B**). This defect was caused by the reduced number of cortical cells (**Fig. 4.28 C**). We also observed that at this point *sdg7-5, sdg24-2* double mutant does not seem to have an additive effect of RAM alterations over the single mutants.

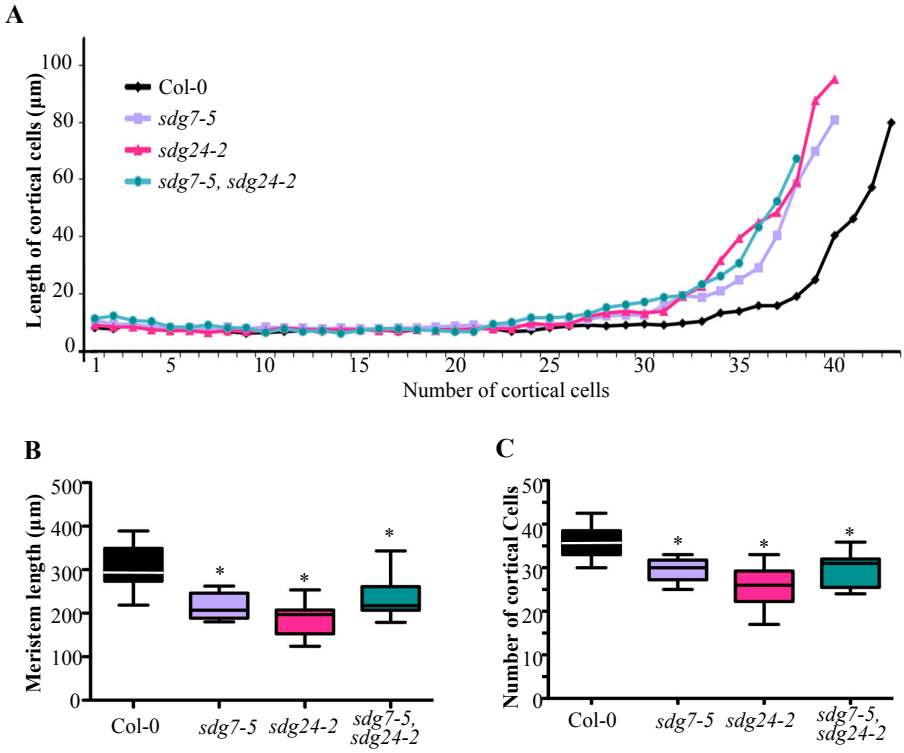


Figure 4.28. Size of the RAM in *sdg7-5* and *sdg24-2* mutants. (**A**) Average size (µm) of root cortical cells along the RAM from the QC (position 1) to the elongation zone. (**B**) Total RAM length distribution. (**C**) Root cortical cell number at the end of RAM. For **A**, **B** and **C** measurements were carried out in 6 dps roots under LD of wild type (Col-0) and mutant lines. Error bars (SD) are shown in panels B and C. Statistical significance compared with Col-0, (*) $P<0.01$, was calculated by the ANOVA-Dunnett's multiple comparisons test, n=10-12 roots. These experiments were done twice, yielding similar results.

Within the RAM a closer look at the SCN also revealed cellular alterations in QC and columella cells (**Fig. 4.29 A**, **B**). Using the specific QC marker, WOX5-GFP, we identified the typical two to three QC cells in the wild type within a single focus plane. In contrast, in *sdg7-5* and *sdg24-2* single mutants those cells presented distorted morphologies, including larger cell size and a disorganized pattern in 50% of *sdg7-5* roots, and 68% for *sdg24-2* roots (**Fig. 4.29 A**, **B**, upper panel). The QC phenotype was usually accompanied by a disarrangement of columella cell layers for both mutants, where three neighboring cells were present instead of four, bilateral symmetry was displaced, or an unusual cells size between detectable (**Fig. 4.29 A, B,** lower panel).

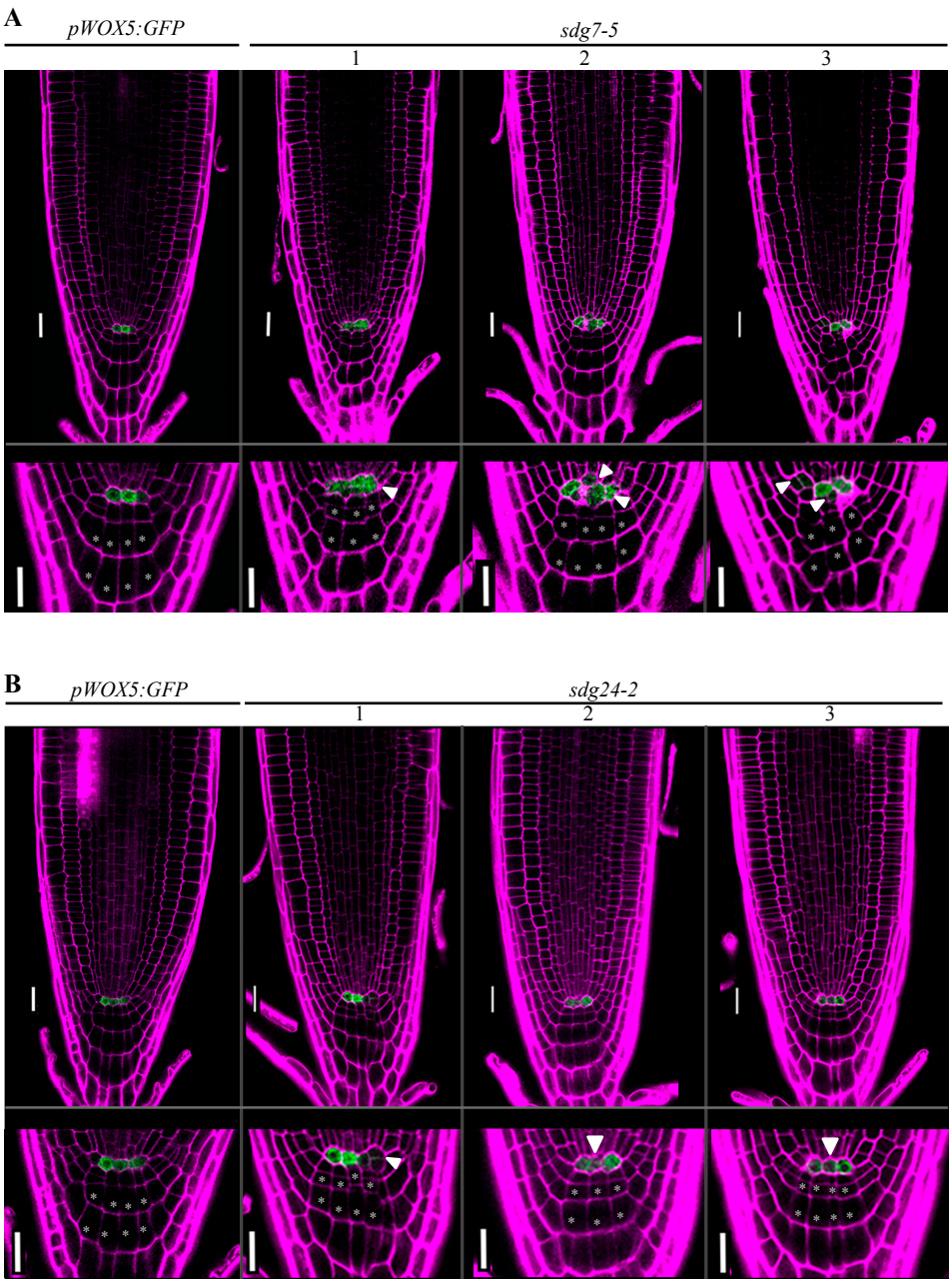


Figure 4.29. **Structure of the stem cell niche and columella cells in *sdg7-5* and *sdg24-2* roots**. Confocal images of 6 dps roots expressing *pWOX5-GFP*. Counterstaining of cell walls was done using propidium iodide (magenta). In each case, the lower panels show an amplified region of the QC and columella cell layers. Arrowheads point to the QC cells. Asterisks (*) denote the 1st and 2nd columella cell layers below the columella initials. Scale bars= 20 μm.

4.2.3.4 Cell divisions are less coordinated in *sdg7* and *sdg24* mutants

The root meristem phenotypes produced by the *sdg7-5* and *sdg24-2* mutations led us to investigate potential alterations in their division capacity. To that end, we monitored cell division events in mutant lines crossed with the pCYCB1;1-GFP reporter line (**Fig. 4.30**). Single planes from QC showed that wild type expression of pCYCB1;1-GFP was predominantly organized in neighboring cells from the same cell file, while in *sdg7-5* and *sdg24-2* expression was confined mainly to several single cells. Maximal projection of images from the QC plane to the epidermis showed that in fact less cells were actively cycling, the *sdg24-2* mutant having the strongest phenotype.

Together, these results suggest that a shorter RAM could be due to a reduced mitotic capacity within the meristematic zone of *sdg7-5* and *sdg24-2* mutants, an effect that was stronger in the *sdg24-2* mutant.

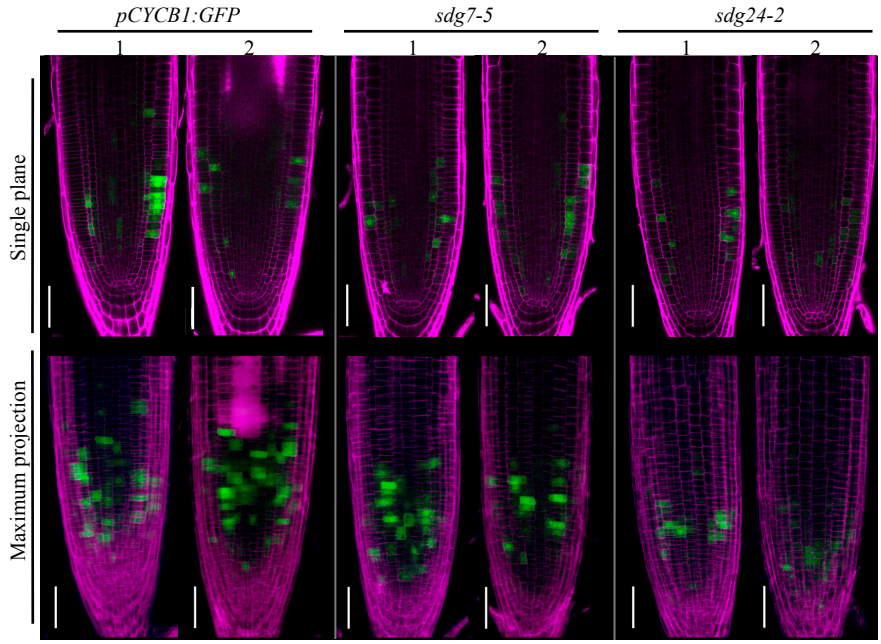


Figure 4.30. Cell division pattern in *sdg7-5* and *sdg24-2* root meristems. Confocal images of 6 dps root meristems of plants expressing *CYCLINB1;1-GFP* (*pCYCB1-GFP*) in the indicated mutant backgrounds. In green, dividing cells in G2-M phase of wild type (Col-0) and mutants lines. Counterstaining was done using propidium iodide (magenta). Midplane of the RAM at the QC and maximum projection images (QC plane to epidermis plane) were generated using the ImajeJ software. Two representative images are displayed for each genotype. We analyzed 20/genotype with similar results. Bars = 50 µm.

4.2.3.5. The S-phase progression is altered in *sdg7* and *sdg24* mutants

To substantiate the aberrant cell division phenotype in our mutant lines, DNA replication was also monitored using a 15 min of EdU pulse treatment. Frist, we confirmed that different root cell types (epidermis, cortex, endodermis and stele) were readily labelled with EdU in the wild type and mutant lines (**Fig. 4.31 A**). Then, we quantified co-replicating

cells, that is, cells that are in S-phase synchronously once they exit the stem cell niche. Thus, daughter, granddaughter, and great-granddaughter of the first replicating cell from the QC will usually result in neighbouring co-replicating duplets (2), quartets (4) and octets (8) as they usually duplicate in pairs. In our analysis we also observed triplets (3) and sextets (6). An illustration of neighbouring replicating cells in the root is depicted in **Fig. 4.31 A**.

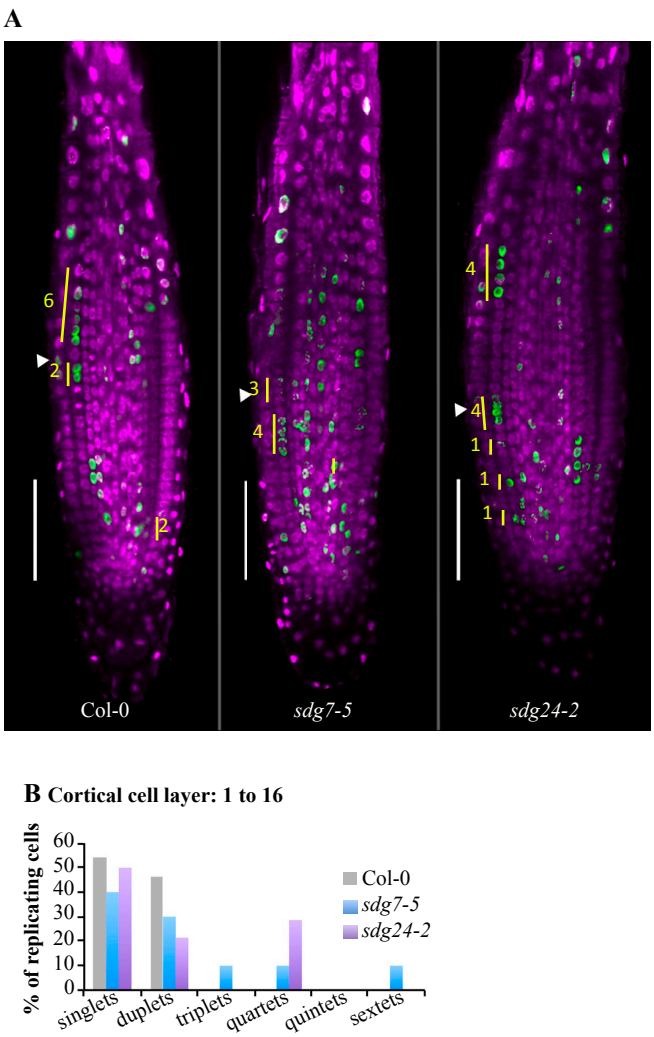


Figure 4.31. DNA replication (S-phase) pattern in roots of *sdg7-5* and *sdg24-2* mutants. (**A**) Representative confocal images of 5 dps primary root meristems of wild type (Col-0) and mutant lines, incubated for 15 min with EdU. In green, EdU-labelled S-phase cells. Yellow lines denote examples of neighbouring replicating cells undergoing synchronously the S-phase: singlets (1), duplets (2), triplets (3), quartets (4) and sextets (6). For counting, meristem was divided in two zones: apical (1-16 cells) and proximal (16-30). Arrowheads denote the position of cell 16th. Counterstaining was done with DAPI (magenta). Scale bars= 50 µm. (**B**) Histogram representing the number of neighbouring replicating cells (%) in the apical part of the RAM. n= 8 roots.

To determine how the stem cells progeny enters in S-phase, and therefore how the cell cycle progresses along the meristem we divided the meristem in 2 zones; the most apical part (cells 1-16), and the proximal part (cells 17-30) counted from the QC to EZ in the cortex. To compare the co-replicating cell pattern we focused on the apical part, where cell proliferation capacity is high. In this area, the wild type presented an accumulation of first stages of co-replicating cell patterns, singlets and duplets, while mutants showed a more wide distribution, between 1, 2, 3, 4 and even 6 co-replicating cells for *sdg7-5* mutant, and 1, 2 and 4 for *sdg24-2* mutant (**Fig. 4.31 B**).

This experiment showed that in both *sdg7-5* and *sdg24-2* mutants there are more cells within the proliferative domain of the RAM replicating DNA, suggesting a direct influence in the root meristem length.

4.2.3.6 rDNA and rRNA patterns are altered in *sdg7* and *sdg24* mutants

In Arabidopsis, the 45S ribosomal gene clusters (NORs) are located at the tip of the short arms of chromosomes 2 and 4 (NOR2 and NOR4) (Pontvianne *et al.*, 2010; Abou-Ellail *et al.*, 2011). In Arabidopsis Col-0, four rDNA variants can be found (VAR1-4). During vegetative growth, variants included in the NOR4 (VAR2, a subset of VAR3 and VAR4) are actively transcribed (rRNA), while variants located within the NOR2 (VAR 1, and the rest of VAR3) will be silenced (Chandrasekhara *et al.*, 2016; Rabanal *et al.*, 2017). This gene dosage control is a system that operates to regulate the number of active rRNA genes according to the physiological needs of the cell (Pikaard, 2003). Chromatin modifications are directly linked to the control of gene activation by generating ON and OFF states in which transcription can proceed (Sequeira-Mendes *et al.*, 2014).

Recent studies have related alterations of rDNA expression due to the defective chromatin modifiers including, among others, HDA6, CHROMATIN ASSEMBLY FACTOR-1 (CAF-1) and the H3K27me1 methyltransferases ATXR5 and ATXR6 (Pontvianne *et al.*, 2010; Pontvianne *et al.*, 2012; Havlova *et al.*, 2016). Taking into account the nuclear expression of SDG7, the nucleolar pattern of SDG24, their capacity for histone binding, and their role in affecting cell proliferation potential, we consider the possibility of finding rDNA defects in *sdg7* and *sdg24* mutants. To do so, we first characterized rDNA gene variants by using genomic DNA from rosette leaves and primers amplifying the 3′ ETS sequence (see Materials and Methods for details). We also included several internal controls, such as the wild type (Col-0) and the *atxr5/6* mutant (**Fig. 4.32**) (Jacob *et al.*, 2009).

Products of semi-quantitative PCR (sqPCR) detected three variants in the wild type: VAR1, VAR2 and VAR3, typically distributed in a ratio of 50-30-20, respectively (**Fig. 4.32 A, B**). In the case of *atxr5/6* mutant, we confirmed that VAR1 was almost the only rDNA variant maintained in the genome (> 80%) as previously described (Pontvianne *et al.*, 2012). The observed pattern in *sdg7-5*, *sdg7-4* and *sdg24-1* was similar to the wild type, with an increased representation of VAR2 (~20% for *sdg7-5*, and ~10% for *sdg7-4 and sdg24-1*). A more striking phenotype was observed in the *sdg24-2* mutant, where VAR1 was almost indistinguishable, and VAR2 contributes over 70% of the rDNA pattern. A similar phenotype

was observed in the *sdg7-5, sdg24-2* double mutant, suggesting that *sdg24-2* mutation is dominant over *sdg7-5*.

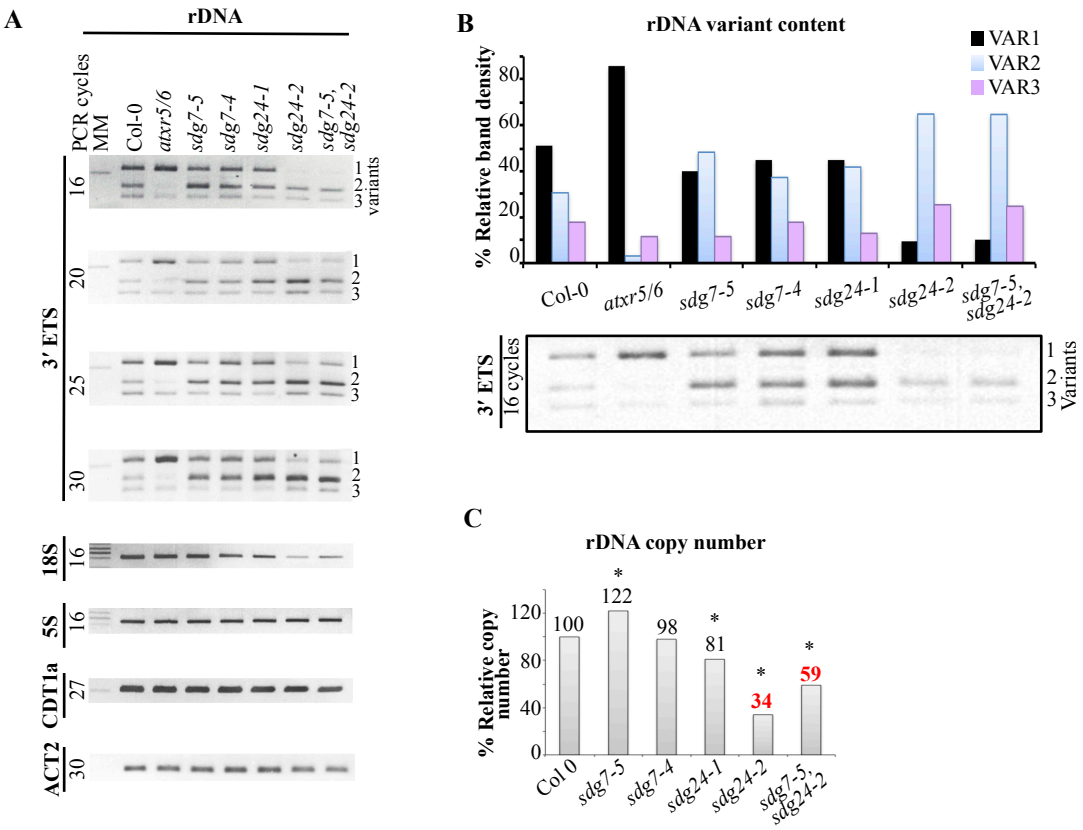


Figure 4.32. rDNA patterns in *sdg7* and *sdg24* mutants. Semi-quantitative PCR detection of rDNA variants in genomic DNA purified from 14 dps rosette leaves of wild type (Col-0) and *atxr5/6, sdg7-5, sdg7-4, sdg24-1, sdg24-2 and sdg7-5/sdg24-2* mutant lines. Different numbers of PCR cycles were performed with all primer pairs and the products were resolved by agarose gel electrophoresis. The *18S* and *5S* were used as multi-copy gene amplification controls, and those of *CDT1a* and *ACT2*, as single-copy gene control. Molecular size of the amplified products of 3'ETS Variant 1, 2 and 3 were: ~460 bp, 430 bp and 410 bp, respectively; for 18S and 5S: ~400 bp; for *CDT1a*: 150 bp, and for *ACT2* ~80 bp. (**B-C**) Relative rDNA variant content in *sdg7* and *sdg24* mutants. (**B**) Relative rDNA patterns measured by quantifying the band intensity of the 3' ETS products. 16 PCR cycles were used and the images were analyzed with the IMAGE-J software. Histogram bars depict variant 1 (VAR1), 2 and 3 in black, blue and magenta, respectively. Normalization was made by dividing individual band intensity by the sum of total variant intensities in each genotype. (**C**) qPCR analysis of total relative rDNA copy number in wild type and mutant lines. The $2^{-\Delta\Delta CT}$ method was used, with *UBQ* as reference. Total percentage (%) of relative copy number is depicted above each bar. (*) $P<0.01$.

The diminished amplification signal of 3′ ETS variants could be directly related to the rDNA gene copy number at NORs (Pontvianne *et al.*, 2013). For this reason we amplified products of a region within the 18S gene by sqPCR (**Fig. 4.32 A**) and estimated the total rDNA copies (in collaboration with Martina Dvořáčková and Jiří Fajkus at the Mendel Centre for Plant Genomics and Proteomics, Brno-Czech Republic) (**Fig. 4.32 C**). In both approaches, the *sdg24-2* mutant and *sdg7-5, sdg24-2* double mutant presented a decreased amplification

pattern, being more evident by qPCR, where *sdg24-2* rDNA copies were estimated to be <60% less than the wild type. Interestingly, we also observed a slight increase of rDNA copies in the *sdg7-5* mutant. This result might be attributed to a real increase of rDNA copies, although no variation was observed by sqPCR (**Fig. 4.32 A**).

Next, we considerer if the number variation was a NOR specific phenomenon or could also occur in other genomic regions, such as the 5S genes, which are located in other chromosomic locations (Cloix *et al.*, 2000). sqPCR analysis of 5S genes and other loci, including *CDT1a* and *ACT2,* showed no visible variations (**Fig. 4.32 A**). These results reinforced the conclusion that the *sdg24-2* mutation causes a loss of rDNA gene copies, preferentially of the loci located at the NOR2.

rDNA analysis in *sdg7* and *sdg24* mutants revealed a first step of gene dosage control, where rDNA copies, detected by 3′ETS variants, are already unbalanced. We hypothesized that rRNA expression pattern may also show alterations in the mutants. To answer this, we analyzed rRNA expression patterns in cDNA from 2-weeks-old rosette leaves. As expected, the wild type revealed that VAR1 was silenced, while VAR2 and VAR3 (of the NOR4) were expressed (**Fig. 4.33 A, B**). Additionally, we confirmed that in the *atxr5/6* mutant VAR1 had the major contribution to the rDNA expression pattern (Pontvianne *et al.*, 2012). In the case of *sdg7-5*, *sdg7-4* and *sdg24-1* mutants, VAR1 was silenced as in the wild type, but VAR2 expression was increased over 20% compared to the wild type, and VAR3 was highly reduced. Interestingly, in the case of the *sdg24-2* and *sdg7-5, sdg24-2* double mutant, the rRNA expression pattern was very similar to the wild type. In addition to the three major variants analyzed, a fourth variant, VAR4, located between VAR1 and VAR2 products, in our experiments was usually faint, and in all cases, it did not showed major differences (**Fig. 4.33**).

In normal conditions silenced and active rDNA copies can be distinguished microscopically. The inactive rDNA clusters are organized in condensed foci outside the nucleolus, while the transcriptionally active parts are in a more decondensed state, with several small foci or uniform patterning within the nucleolus (Dvorackova *et al.*, 2015). With the previous results, we speculated that *sdg24-2* would contain fewer foci around the nucleolus, as the inactive copies have been dramatically reduced. To test this hypothesis, a fluorescent *in situ* hybridization (FISH) analysis of rDNA was carried out in collaboration with the Fajkus laboratory. The nuclei from wild type roots showed the typical pattern described above, where two brighter rDNA foci were observed at the periphery of the nucleolus, while some dispersed signal was located inside of it (**Fig. 4.34**).

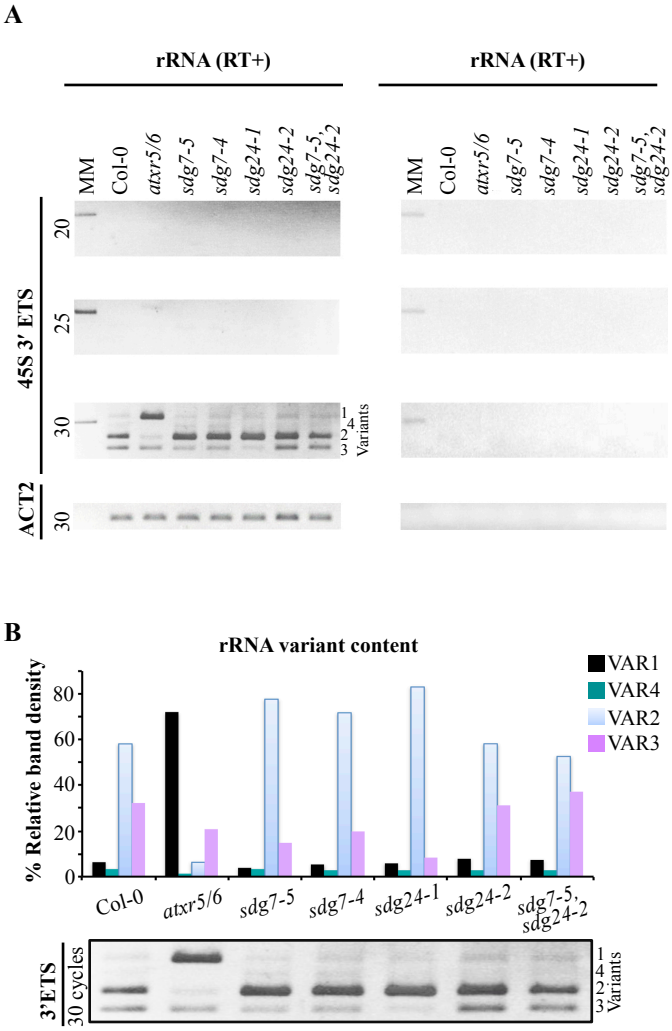


Figure 4.33. Expression pattern of rRNA variants in *sdg7* and *sdg24* mutants. (A) Semi-quantitative PCR detection of rRNA variants in reverse transcribed (RT+) cDNA purified from 14 dps rosette leaves of wild type (Col-0) and mutant lines. Different number of PCR cycles were performed with all primer pairs, and the products were resolved by agarose gel electrophoresis. *ACT2* served as positive control. Reactions lacking reverse transcriptase (RT-) served as negative controls. **(B)** Relative rRNA expression was quantified by measuring band intensities of 3′ETS variants with the IMAGE-J software (reactions carried out for 30 PCR cycles). Colored bars depict variant 1 (VAR1), 2, 3 and 4 in black, blue, magenta and green, respectively. Normalization was made by dividing individual band intensity by the sum of total variant intensities in each genotype.

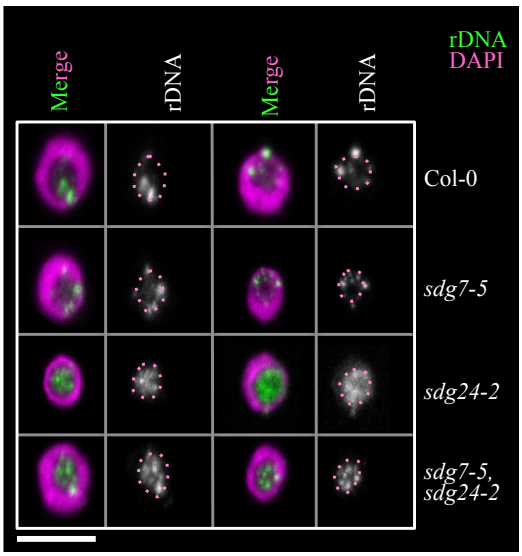


Figure 4.34. Analysis of 45S rDNA in *sdg7* and *sdg24* mutants by fluorescent *in situ* hybridization (Ikura *et al.*). Nuclei from the root tip of the wild type (Col-0) and mutant lines were hybridised with a 45S rDNA probe (green) and counterstained with DAPI (magenta). rDNA channel is presented apart in grey where the nuclear area is delimited. Two confocal images are shown per genotype. Bar= 10 µM.

This analysis also included the *sdg7-5* mutant, which was resembled to the wild type, although with smaller rDNA foci surrounding the nucleolar area and some decondensation inside the nucleolus. Interestingly, the *sdg24-2* nuclei presented a different scenario, where most of the rDNA signal colocalized inside the nucleolus in a more decondensed manner. Similarly, the *sdg7/sdg24-2* mutant presented a comparable pattern with *sdg24-2*. With this approach we confirm our initial data obtained for *sdg24-2*, where in fact there is a loss of inactive rDNA variants, and the remaining active copies are inside the nucleolus. However, the presence of more foci around the nucleoli and decondensation in the case of *sdg7-5* it might indicate in fact more rDNA copies.

Since rDNA expression might vary according to the plant developmental stage (Pontvianne *et al.*, 2010), we also addressed the rDNA and rRNA patterns in different tissues of the plant, including: seeds, roots, cotyledons, vegetative SAM and flowers (**Fig. 4.35**). At the DNA level all tissues were very similar between them, as was described above in the rDNA analysis for *sdg7* and *sdg24* mutants (**Fig. 4.35**, left panels). However we found some interesting variations at the expression level when compared to the wild type, in particular during early vegetative development, e.g. imbibed seeds (**Fig. 4.35**, middle panels). As expected, in the wild type all variants were expressed in the 24 hour imbibed seeds (Pontvianne *et al.*, 2010), however the *sdg7-4, sdg7-5* and *sdg24-1* mutants showed an increased expression of VAR1 and VAR2 compared to the wild type. Interestingly, the *sdg24-2* and the *sdg7-5, sdg24-2* double mutant did also show some VAR1 expression, while the VAR2 was highly expressed. These results might indicate an altered expression mechanism in the *sdg7* and *sdg24* mutants at early stages of development.

In general, our rDNA analysis showed that *sdg7* and *sdg24* mutations produced alterations in the rDNA expression pattern. In particular, the lost of rDNA gene copies in the *sdg24-2* mutant seem to be altering the entire NOR2, as VAR1 associated to this region is the most affected. Interestingly, rRNA analysis in *sdg24-2* showed that the loss of inactive variants is compensated with the continuous expression of active variants (VAR2 and VAR3) from early stages of development, e.g. seeds, since VAR1 was not expressed. Taken together these results, we suggest that both SDG7 and SDG24 are involved in controlling rDNA copies in the two NORs and also in modulating rRNA expression inside the nucleolus.

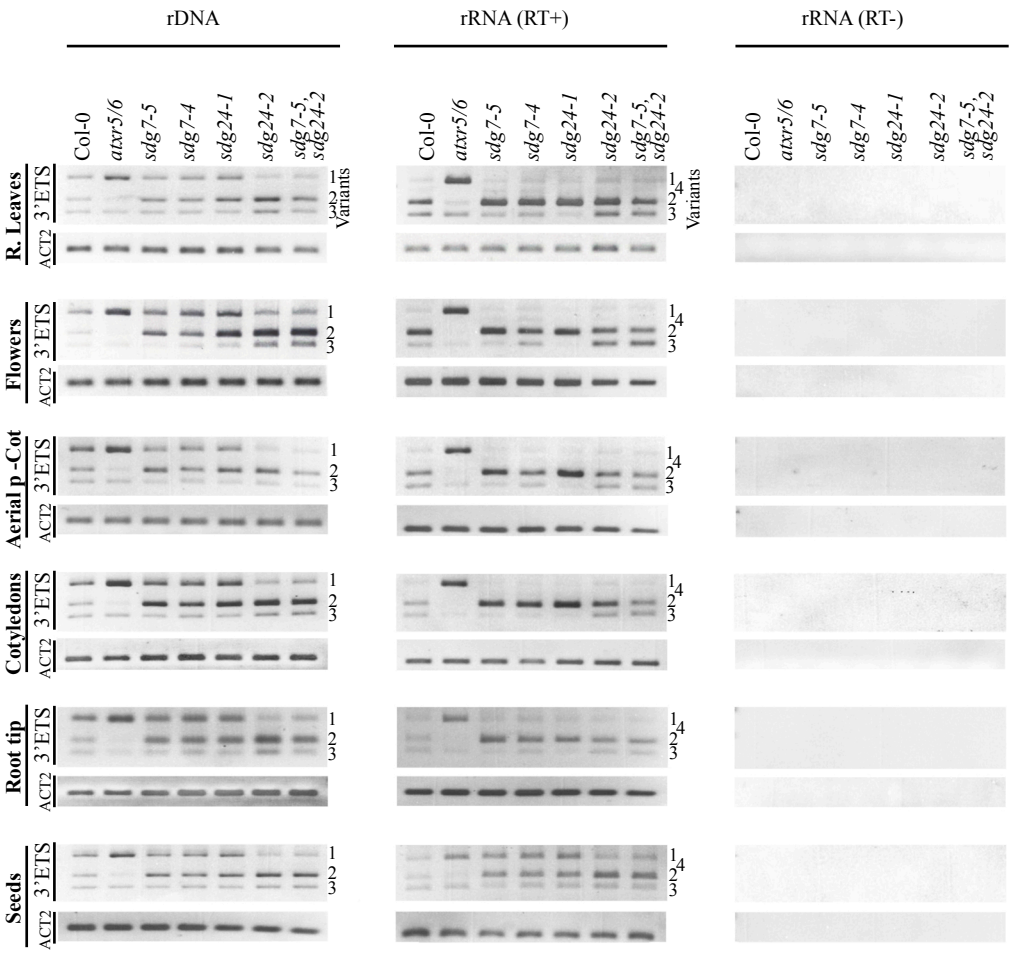


Figure 4.35 rDNA variants and rRNA expression patterns in different developmental stages of Arabidopsis. Analysis of rDNA variants and rRNA expression in wild type (Col-0) and mutant lines from different Arabidopsis tissues. Extractions were carried out from 14 dps rosette leaves (R. leaves), full developed flower buds, 7 dps cotyledons, 7 dps vegetative SAM (aerial part except cotyledons, Aerial p- Cot), 7 dps root tips (0.5 cm), and seeds imbibed 24 h in H_2O. PCR cycles were applied (25 and 30) for 3′ ETS primers in rDNA and rRNA (RT+ and RT-), respectively. *ACT2* was used as positive control, with 30 cycles of amplification. Reactions lacking reverse transcriptase (RT-) served as negative controls.

4.2.3.7 Transgenerational maintenance of the rDNA pattern in *sdg7* and s*dg24* mutants

After characterizing rDNA variations in the *sdg7* and *sdg24* mutants, we asked for the transgenerational maintenance of the pattern observed. To address this question we backcrossed *sdg7-5* and *sdg24-2* mutants with the wild type and analyzed the 3′ETS variant distribution in 2 weeks-old rosette leaves (**Fig. 4.36**). Already since generation 1 (F1), heterozygous plants displayed unbalanced distributions of variants in both mutants, establishing the typical rDNA pattern at the F3 (second homozygous generation) for *sdg7-5*, and at F2 (first homozygous generation) for *sdg24-2*. This points to early mechanisms interfering with the normal maintenance of the rDNA reservoir.

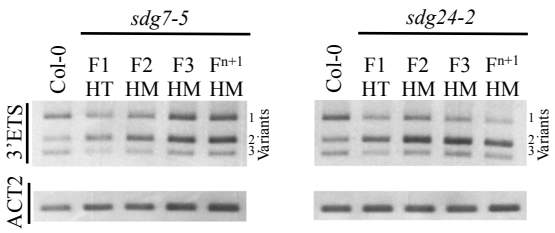


Figure 4.36. Pattern of rDNA variants in different successive generations of *sdg7* and sdg24 mutants. rDNA pattern analysis in back-crossed *sdg7-5* and *sdg24-2* mutants lines with wild type (Col-0). Genotyping of each mutation was performed in heterozygous generation 1 (F1-HT) and homozygous (HM) F2, F3 and F^{n+1}. Genomic DNA was extracted from 14 dps rosette leaves. PCR cycles (25 and 30) were performed for 3′ ETS and *ACT2*, respectively.

We also tested if the rDNA pattern changed in nuclei of different ploidy levels (**Fig. 4.37**). To do so, we FACS-sorted nuclei from 3 weeks-old rosette leaves and extracted gDNA from cell populations containing 2C, 4C, 8C and 16C. Results from sqPCR analysis of 3′ETS showed that all cell populations shared the same rDNA pattern observed previously for Col-0, *atxr5/6*, and for the mutants *sdg7-5* and *sdg24-2*, indicating that the rDNA pattern is maintained during endoreplication.

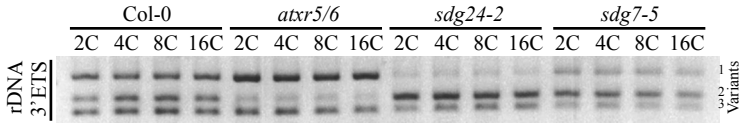


Figure 4.37 Pattern of rDNA variants in nuclei with different ploidy levels. rDNA pattern analysis of FACS-sorted nuclei with different ploidy levels (2C, 4C, 8C and 16C). wild type (Col-0) and mutant lines *atxr5/6, sdg24-2 and sdg7-5* were analyzed. Genomic DNA was isolated and quantified with a Qubit fluorimeter. Equal amounts of DNA were used as substrate for PCR amplification using 3′ ETS primers and 30 cycles.

4.2.3.8 Telomere length in *sdg7* and *sdg24* mutants

The rDNA repeats of NOR2 and NOR4 are capped by telomere repeats that are added directly to rDNA gene sequences (Copenhaver and Pikaard, 1996). In mutants of some chromatin remodelers, e.g. CAF-1 (*fas1* and *fas2*), loss of rDNA copies was also accompanied by a decreased in telomere repeated sequences (Mozgova *et al.*, 2010). From our previous results of rDNA copy loss in the *sdg24-2* mutant, we evaluated the possibility of an altered telomere size in this line, again in collaboration with the Fajkus laboratory. Telomere length was measured by a terminal restriction fragment analysis (TRF) on genomic DNA from the wild type and mutant lines (**Fig. 4.38**). Preliminary results of TFR in the wild

type Col-0 (lines 1-3), and the mutants *sdg7-5* (lines 7-9) and *atxr5/6* (lines 13-15), showed that telomeres ranged in size from 2 to 4 kb, indicating that no significant changes in telomere length occurred. However, telomere length of *sdg24-2* mutant was shorter, spanning from 1.5 to 3.5 kb (lines 10-12). Some telomere shortening was also observed for *sdg7-5, sdg24-2* double mutant (lines 4-6), although not as clear as in *sdg24-2*. Together these results suggest that in addition to rDNA copy loss, *sdg24-2* mutant also possesses shorter telomeres, may be associated to altered capping of the NOR2 repeats.

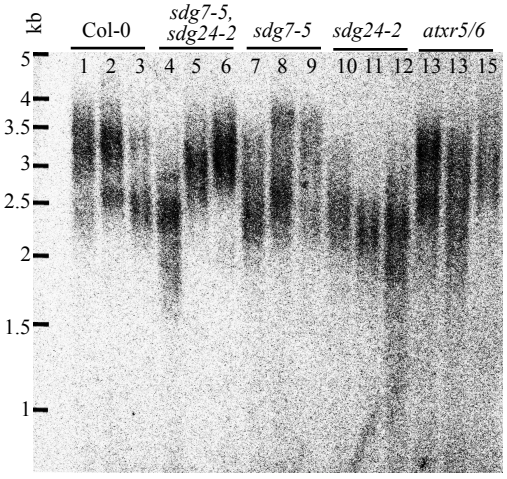


Figure 4.38. Telomere length in *sdg7* and *sdg24* mutants. Genomic DNA of the wild type (Col-0) and mutant lines were subjected to a terminal restriction fragment analysis (TRF). Products were detected by using the telomeric TR4C end-labelled with [γ-^{32}P]ATP. Molecular markers (kb) are shown in the left of the image.

4.2.3.9 DNA damage response in *sdg7* and *sdg24* is not dramatically altered

As a consequence a decreased rDNA gene copy number in *sdg24-2* mutant we wondered if DNA damage response was altered as it occurred in the *atxr5/6* mutant (Feng *et al.*, 2017). By measuring the expression level of *BREAST CANCER SUSCEPTIBILITY-1* (*BRCA1*), *POLY (ADP-RIBOSE) POLYMERASE-1* (*PARP1*) and the DNA-dependent ATPase *RAD51* we can detect if DNA repair pathways are activated. *BRCA1* and *RAD51* are particularly involved in double-strand break (DSBs) repair by the homologous recombination (HR) pathway, and changes in their expression levels are indicative of DNA damage occurrence (**Fig. 4.39**) (West *et al.*, 2004; Seeliger *et al.*, 2012; Song *et al.*, 2015). We used the *atxr5/6* mutant as positive control since, its heterochromatic overreplication phenotype leads to a high expression of these checkpoint genes (Feng *et al.*, 2017). qPCR analysis in *sdg7-5*, *sdg24-2* and *sdg7-5, sdg24-2* double mutant showed that *BRCA1* and *RAD51* had a lightly increased expression compared to the wild type, but not comparable to the fold change observed in *atxr5/6*.

Together, with these results we speculate different scenarios. One would be that the DNA damage response is not altered in *sdg7-5* and *sdg24-2* mutants due to DSBs, and

changes in rDNA copy number might result from an alternative mechanism of HR. Other possibility could involve DSB response acting in a constricted area, localized perhaps only at the NORs; or by the contrary, expression of *RAD51* and *BRCA1* at this levels, could indicate only basal level of expression.

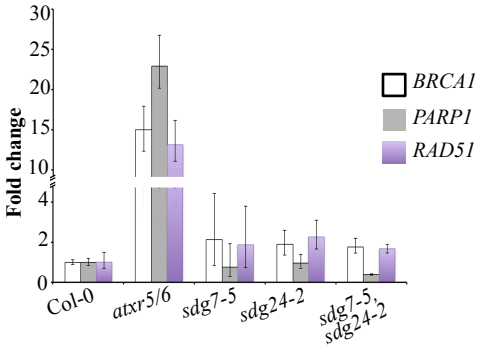


Figure 4.39. Expression of DNA damage response genes in *sdg7-5* and *sdg24-2* mutants. qPCR analysis of *BRCA1*, *PARP1* and *RAD51* genes in various genotypes. The relative expression of these marker genes in wild type (Col-0) and mutant lines was obtained by measuring the fold change ($2^{-\Delta\Delta CT}$ method) in 14 dps rosette leaves. Error bars correspond to the SD between technical replicates. *GADPH2* served as gene reference. Root tips were also tested, showing similar expression patterns (not shown).

4.2.3.10 Nucleolar dominance of *sdg7* and *sdg24* with other lysine methyltransferases

One manifestation of gene dosage control in hybrids is nucleolar dominance, an epigenetic phenomenon in which rDNA genes of one progenitor are repressed (Tucker *et al.*, 2010). In Arabidopsis an analogous nucleolar dominance-like phenomenon also occurs in non-hybrids, in which specific classes of rDNA gene variants are selectively inactivated during early vegetative development (Pontvianne *et al.*, 2012). This phenomenon can be observed in the *atxr5/6* double mutant, where VAR1, instead of VAR2-3, is the active variant (Pontvianne *et al.*, 2012). Moreover, a cross amongst different HMTase mutants, including the quintuple mutant *atxr5/6, suvh4/5/6* (SUVH4-5; H3K9me1 HMTases), induces wild type-like arrangements of rDNA and rRNA expression (Pontvianne *et al.*, 2012). This indicates that altering different histone marks creates different nucleolar expression scenarios. In the present work we also spotted this phenomenon in the *sdg7-5, sdg24-2* double mutant, where the *sdg24-2* mutation was dominant over *sdg7-5*, letting us to speculate that *sdg24-2* could interfere with other HMTases to regulate rDNA maintenance.

Based on the peptide-binding assay results, where we found that both SDG7 and SDG24.1 had some preferential recognition for H3K27 marks, and that ATXR5/6 (the HMTase that deposits H3K27me1) have an altered rDNA pattern of the VAR1, we speculated for a possible interaction between them. Therefore, we generated the triple mutants *atxr5/6, sdg7-5* and *atxr5/6, sdg24-2,* in both cases *atxr5/6* acting as the maternal donor. In our first observations of plant growth (under LD conditions) became evident that both triple mutant

plants had an apparent retarded growth and less fruit production, compared to the wild type and parental lines (**Fig. 4.40 A, B**). Remarkably, roots from the *atxr5/6, sdg24-2* triple mutant seem to have shorter LRs (**Fig. 4.40 C**).

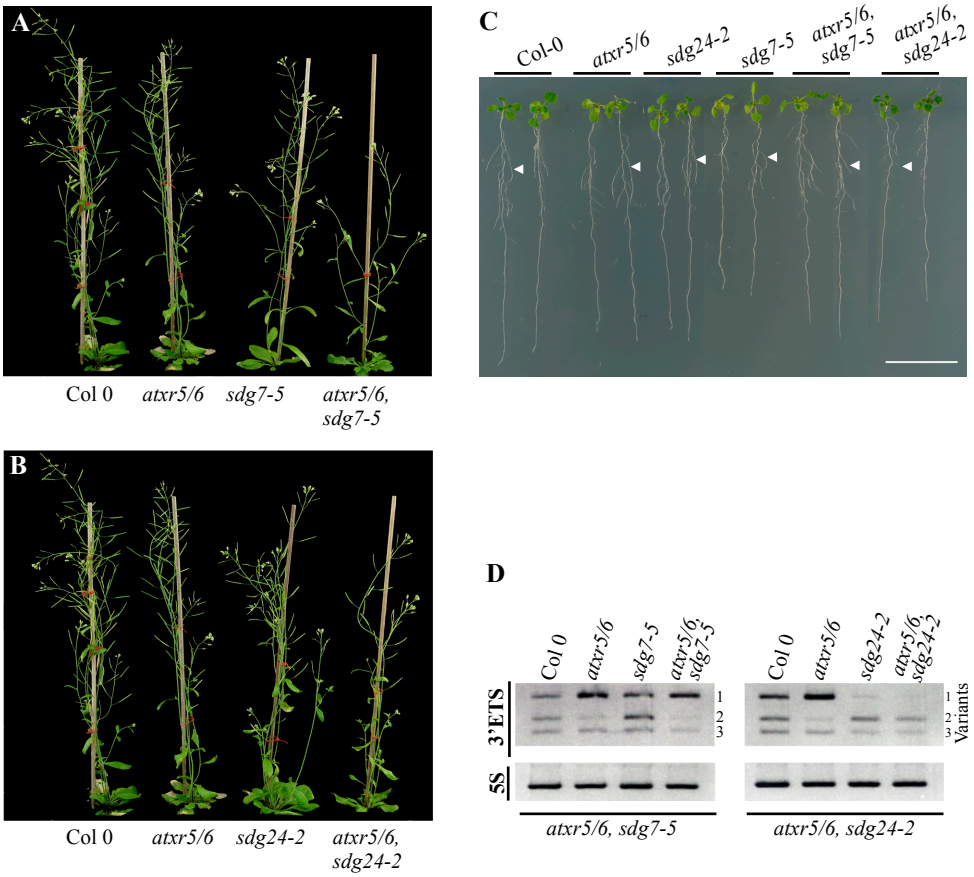


Figure 4.40. Interaction of SDG7 and SDG24 with the ATXR5/6 H3K27me1 HMTases. Growth phenotype of the *atxr5/6, sdg7-5* (**A**) and *atxr5/6, sdg24-2* (**B**) triple mutant plants in 6 weeks-old plants from wild type (Col-0) and mutant lines grown under LD conditions. (**C**) Root growth of the triple mutants *atxr5/6, sdg7-5*, and *atxr5/6, sdg24-2,* and their respective parental lines. Arrowheads point to lateral roots. Scale bar = 2 cm. (**D**) Relative abundance of rDNA variants in *atxr5/6, sdg7-5* and *atxr5/6, sdg24-2* triple mutants. Genomic DNA from 21 dps rosette leaves was used as template to measure rDNA copies. The *5S* panel shows that equal amounts of DNA were subjected to PCR for each genotype. 16 and 20 PCR cycles were used to identify the 3'ETS and *5S* regions, respectively. Analysis of triple mutants was conducted with plants of the F2 homozygous generation, using the *atxr5/6* double mutant as maternal donor.

Next, we determined whether the *sdg7-5* or *sdg24-2* mutations could alter the rDNA pattern of *atxr5/6*. Analysis of *atxr5/6, sdg7-5* triple mutant showed a clear *atxr5/6* pattern, where VAR1 remains more abundant than VAR2 and VAR3 (**Fig. 4.40 D**, left side of the panel). On the contrary, the *atxr5/6, sdg24-2* triple mutant (right side of the panel**)** displayed the typical *sdg24-2* phenotype, where VAR1 was almost undetectable. These results reveal that the *sdg24-2* mutation may act upstream of *atxr5/6.* Interestingly, dominant patterns of *atxr5/6* over *sdg7-5*, and *sdg24-2* over *atxr5/6* were already visualized in different

homozygous-heterozygous segregating lines of the parental lines (not shown). We also noticed that *sdg24-2* imprinted a dominant pattern despite its parental contribution, mother o father, as the same rDNA phenotypic display was established when *sdg24-2* was used in the *sdg7-5, sdg24-2* double mutant and in the *atxr5/6, sdg24-2* triple mutant.

Our findings suggest that SDG24 activity is controlling the stability of NORs, possibly by interacting with the silencing mark H3K27me1 at the NOR2 mediated by ATXR5/6 proteins, and resulting in the lost of inactive rDNA copies of NOR2. Also, the unaffected *atxr5/6* pattern by the *sdg7-5* mutation suggests that SDG7 activity is downstream of SDG24 and ATXR5/6. Together these results led us to hypothesize that SDG24 might regulate ATXR5/6 function particularly at NORs.

Discussion

5.1 SDG7: a nuclear and cytoplasmic protein

The cellular approach undertaken in this study has shown that SDG7 protein possesses a nuclear and cytoplasmic localization in root cells. Our findings are in agreements with the presence of putative NLS signals identified with several bioinformatic tools, including the cNLS Mapper, seqNLS and NLStradamus. This is in contrast with previous studies reporting the absence of any NLS signals along the SDG7 sequence (Sahr *et al.*, 2010; Lee *et al.*, 2015), where predictions were made with alternative software (WoLF PSORT, TargetP, Predotar and NucPred). Based on this information, results published by Sahr et al. (2010), using the CaMV 35S promoter sequence driving *SDG7* expression, concluded that SDG7 localized to the membranes of root epidermal cells, in the elongation-differentiation zone. In contrast, we used a ~2 kb upstream the start codon of *SDG7* as the endogenous promoter region. It is known that promoters regulate gene expression through DNA recognition sequences, which interact with basic transcription initiation complexes and numerous transcription factors (Hernandez-Garcia and Finer, 2014). Gene expression in plants have shown that the majority of elements necessary for an important regulation of expression lie immediately upstream of the transcriptional initiation site (TSS), usually within 1-2 kb region (Hernandez-Garcia and Finer, 2014). Although the use of constitutive promoters had been widely implemented in plant research, they may provide ectopic gene expression in transgenic plants that might not reflect truthfully protein localization under normal conditions (Dutt *et al.*, 2014). Thus it is possible that the use of different promoters and perhaps different protein expression levels are responsible for the differences observed.

Our results reinforce the notion that some available bioinformatic tools and algorithms might be more robust than others in predicting the subcellular localization of proteins. In addition, we noticed that optimal confocal microscopic conditions could be a limiting factor regarding *in vivo* localization of proteins with a low expression level, as this was a key factor to discern SDG7 inside the nuclei.

5.2 Alternative splicing of *SDG24* results in different expression patterns

At the start of this project *SDG24* was a not predicted to produce alternatively spliced variants. However our initial transcriptional analysis confirmed the presence of several transcripts coexisting in flowers. When the Arabidopsis Information Portal (araport.org) was updated, *SDG24.1* and *SDG24.2* (*AT3G59960.2* and *AT3G59960.1*, respectively) were reported as the main transcript contributors of four different versions, but no reference was found for the *SDG24.3* and *SDG24.4* variants described in this work. Interestingly, the two

other splicing variants presented in Araport (*AT3G59960.3* and *AT3G59960.4*) were similar to the *SDG24.1* coding sequence, but differed in their 3′ untranscribed region (**Fig. 5.1**). Alternative splicing in genes occurs through various mechanisms (Nilsen and Graveley, 2010). From our results, it seems that in *SDG24* some of them are employed, including alternative donor (AltD), alternative acceptor (AltA) and exon skipping (ExonS) (**Fig. 5.1**).

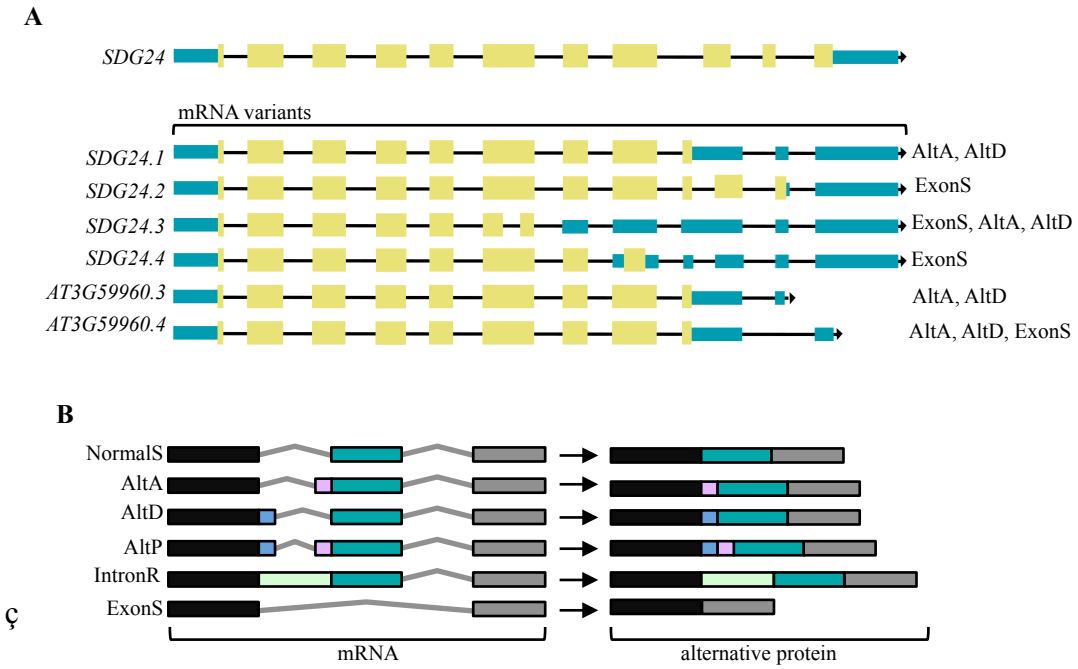


Figure 5.1. Alternative splicing patterns. (**A**) SDG24 mRNA splicing variants encode different ORF. At the left side of each diagram are described types of splicing in SDG24 mRNA transcripts. Yellow boxes: translated region; Green musk: untranslated region. Black line: intronic sequences. (**B**) Model of alternative splicing patterns. The left side of the diagram shows normal and five major alternative splicing patterns for a representative gene containing three exons (black, green musk and grey boxes, respectively). Pink, blue and pale green boxes denote alternative exonic sequences. The edited exons after splicing are shown to the right. NormalS, normal splicing; AltA, alternative acceptor (3' side of introns), AltS, alternative donor (5' side of introns); AltP, alternative positions (both 5′ and 3′ side of introns); IntronR, intron retention; ExonS, exon skipping. Grey lines: intron regions. Modified from Ng et al., 2006; Nilsen and Graveley, 2010.

As mentioned earlier, the majority of elements that modulate gene expression lie within the 5′ region upstream of the TSS (Hernandez-Garcia and Finer, 2014); however, in some cases these regulatory signals may be located within introns or at the 3′UTR (Larkin *et al.*, 1993; Raatz *et al.*, 2011; Serrano-Mislata *et al.*, 2016). Alternative splicing in non-translated regions can have an impact on protein expression levels and localization, while in coding regions it may alter protein structure and function, and even lead to the production of proteins with unrelated functions (Ng *et al.*, 2007; Hernandez-Garcia and Finer, 2014). For example, the Drosophila SET domain *Su(var)3-9* gene can express two distinct transcripts (2.4 kb and 2.0 kb) encoding two different proteins that share identity only for the first 80 amino acids: one of them yields a SET domain-containing HMTase, whereas the other splicing variant

encodes the γ subunit of eukaryotic TRANSLATION INITIATION FACTOR-2 (γeIF2) (Krauss and Reuter, 2000). The presence of several SDG24 variants, including one of them with an altered SET domain, supports the hypothesis of different proteins with diverse functions being produced from the *SDG24* gene.

In Arabidopsis, alternative splicing of SET domain coding sequences is a common feature. The updated version of Araport reveals that >60% of them show some form of alternative splicing (**Table 5.1**). This proportion is substantially higher than the average fraction (42%) for Arabidopsis genes (Filichkin *et al.*, 2010), and suggests that alternative splicing is important in generating SET domain protein complexity and functionality. For instance, an intron located in the 5′UTR of *SUVH3* regulates both constitutive and high levels of expression (Casas-Mollano *et al.*, 2006), while alternative splicing in *SUVR1* and *SUVR4* generated proteins with different subcelular localization, in nucleus and nucleolus (**Table 5.1**) (Thorstensen *et al.*, 2011).

In this work we found that SDG24.1 was particularly expressed in nucleoli from root cells, which utterly led us to identify a role of SDG24 in the rDNA maintenance. Interestingly, a recent study in *Brassica rapa,* a close relative of Arabidopsis, its SDG24 homologue, the BrASHH4a, exhibited a location in both the nucleus and cytoplasm of root cells (Dong *et al.*, 2015), which could imply a partial functional homology between this two related species.

Expression of alternatively spliced transcripts can be also tissue specific (Hernandez-Garcia and Finer, 2014). Splicing in *SDG24* resulted in different spatial expression patterns, being the *SDG24.1* specifically expressed in root cells*,* and to the nucleus of the central cell in the mature embryo sac. Interestingly, two Arabidopsis E(Z) homologs, SWN and MEA, required for the maintenance of H3K27me3 levels, are also expressed in the central cell, and together perform a partially redundant role in controlling the initiation of endosperm development before fertilization in Arabidopsis (Wang *et al.*, 2006). Expression of SDG24 at this particular cell, prior pollination, might control molecular events related to the maturation process of the female gametophyte, and perhaps in relation to the rDNA maintenance and or expression.

Table 5.1. Splicing variants in Arabidopsis SET domain proteins

Class	Name	SDG number	Number of splicing	Splicing Outcomes	Subnuclear localization	Ref
I- E(Z)	MEA	SDG5	2	AltD: Alternative stops	Nucleus and cytoplasm (central cell of ovule)	Wang et al., 2006
	CLF	SDG1	2	ExonS: Alternative starts	Nuclear Absent in in mitosis	Schubert et al., 2006
	SWN	SDG10	4	IntronR, AltD: Alternative exons at 3´	Nuclear (central cell of ovules, trichomes, root, LR)	Wang et al., 2006
II- ASH1	ASHH1	SDG26	4	IntronR: Alternative 5´UTR	Nuclear	Xu et al., 2008
	ASHH2/ ESF	SDG8	2	IntronR, ExonS: Alternative exons at 3´	Nuclear	Xu et al., 2008
	ASHR3	SDG4	None described	---	Nuclear (euchromatin)	Thorstensen et al., 2008
	ASHH3	SDG7	None described	---	ER Nuclear	Sahr et al., 2010 This Work
	ASHH4	SDG24	6	IntronR, AltD, ExonS: Alternative exons Alternative stops	Nucleolar (root) Nuclear (central cell of ovules, trichomes)	This Work
III- TRX	ATX1	SDG27	None described	None described	Nuclear	Saleh et al., 2007
	ATX2	SDG30	4	IntronR: Alternative 3´UTR	Inferred nuclear pattern	
	ATX3	SDG14	2	IntronR: Alternative exons at 3´	Inferred nuclear pattern	
	ATX4	SDG16	None described	---	Inferred nuclear pattern	
	ATX5	SDG29	None described	---	Inferred nuclear pattern	
	ATXR3	SDG2	None described	---	Inferred nuclear pattern	
	ATXR7	SDG25	8	IntronR, ExonS: Alternative 5´- 3´ UTR	Nuclear	Berr et al., 2009
IV-	ATXR5	SDG15	2	3´UTR	Nuclear Chloroplast	Raynaud et al., 2006
	ATXR6	SDG34	None described	---	Nuclear	Raynaud et al., 2006
V- SU(VAR) 3-9	SUVH1	SDG32	2	AltD: Alternative 3´UTR	Nuclear /Subnuclear foci	Naumann et al., 2005
	SUVH2	SDG3	4	ExonS: Alternative 5´UTR	Nuclear /Subnuclear foci	Naumann et al., 2005
	SUVH3	SDG19	None described	---	Nuclear /Subnuclear foci Mitotic chromosomes	Casas-Mollano et al., 2006
	SUVH4/ KYP	SDG33	2	IntronR: Alternative starts	Nuclear	Yu et al., 2017
	SUVH5	SDG9	5	IntronR: Alternative 5´ UTR, Alternative starts	Nuclear	Yu et al., 2017
	SUVH6	SDG23	2	AltD, ExonS: Alternative 5´-3´UTR	Nuclear	Yu et al., 2017
	SUVH7	SDG17	None described	---	---	
	SUVH8	SDG21	None described	---	---	
	SUVH9	SDG22	2	IntronR, AltD: Alternative 5´-3´UTR		
	SUVH10	SDG11	None described	---	---	
	SUVR1	SDG13	3	ExonS, IntronR: Alternative stars, Alternative 3´UTR	Nucleolar Nuclear (foci)	Thorstensen et al., 2006 Han et al., 2014
	SUVR2	SDG18	17	AltD, AltA, IntronR: Alternative 3´UTR Alternative start Alternative exons	Nuclear (foci)	Thorstensen et al., 2006 Han et al., 2014
	SUVR3	SDG20	2	AltD, AltA Alternative 5´-3´UTR Alternative exons	Inferred nuclear pattern	Ng et al., 2007
	SUVR4	SDG31	6	AltD, IntronR: Alternative 5´-3´ UTR, Alternative start, stop	Nucleolar	Thorstensen et al., 2006
	SUVR5	SDG9	6	AltD, ExonS: Alternative 5´-3´UTR Alternative stop	Nuclear	Krichevsky et al., 2007

AltD; Alternative starts: different Star codon; alternative stops: different stop codons. AltA: alternative acceptor (3' side of introns), AltS: alternative donor (5' side of introns); IntronR: intron retention; ExonS: exon skipping.

5.3 SDG7 and SDG24 might recognize different chromatin scenarios

SDG7/ASHH3 and SDG24/ASHH4 are classified as Arabidopsis class II SET domain group proteins, based on their overall structure resemblance to Drosophila ASH1. In our search for H4K20 HMTases, we investigated whether loss of *SDG7* and *SDG24* affected the H4K20me3 mark by Western blot. However, none of the tested mutants resulted in

modification of this mark, although we observed a high lack of specificity of antibodies to detect both H4K20me1 and H4K20me3 marks. In addition, we considered alternative histone residues as substrates for both proteins; however, the *in vitro* HMTase assay did not show alteration of H4K20 and others marks, including H3K27me1/me2/me3, H3K4me2/me3, H3K9me2/me3 and H3K36me3. Interestingly, during the course of this project, similar negative results were obtained for SDG7 by Lee et al. (2015), in which they tested the *in vitro* activity with the incorporation of ^{14}C-labeled SAM into oligonucleosomes, mononucleosomes, and free histones. Despite these negative results, SDG7 was found to be required for proper induction of the VERNALIZATION INSENSITIVE 3 (VIN3), a cold-specific PRC2 component. A common feature of the Lee et al. study and our work is that the expression of SDG7 was done in a prokaryotic system. It is possible that some PTMs in the protein are needed in order to display enzymatic activity. For instance this is the case of SUVR4 protein, in which the binding to ubiquitin facilitates the addition of two methyl groups to the H3K9me1 (Veiseth *et al.*, 2011; Rahman *et al.*, 2014).

The collective negative results of SDG7 on its methyltransferase activity on histones are particularly intriguing, as this protein was initially described to posses methyltransferase activity on the C-terminal side of the PIP2;1 protein, an Arabidopsis aquaporin of the plasma membrane (Sahr *et al.*, 2010). Other HMTs, including G9A (H3K9), PR-SET7/SETD8 (H4K20) and SUV4-20H (H4K20), have been reported to target non-histone proteins (Ying *et al.*, 1996; Herz *et al.*, 2013; Milite *et al.*, 2016; Weirich *et al.*, 2016). In particular, the Arabidopsis SUVH2/SUVH9 SET domain proteins lack *in vitro* HMTase activity, but can bind methylated DNA (cytosines) through their SET and RING-associated (SAR) domains. Together, SUVH2/SUVH9 mediate Pol V recruitment in RNA-directed DNA methylation pathway (RdDM), and play an important role in the recruitment of chromatin-remodeling SWI/SNF complex to target loci by associating with the Microrchidia (MORC) adenosine triphosphatase proteins (Kuhlmann and Mette, 2012; Liu *et al.*, 2016; Jing *et al.*, 2016). A similar mechanism may be used by SDG7 to impinge on rDNA stability and –or– expression.

Some SET domain-containing proteins are found in complexes or interacting with proteins that regulate their target specificity and catalysis (Herz *et al.*, 2013). In Arabidopsis, yeast two hybrid assays have demonstrated also interactions between ASHH1/ASHH2 and ASHH1/ATX1 proteins (Valencia-Morales *et al.*, 2012). Interestingly, ATX1 has been shown to be required for H3K4me3 deposition at the *AGAMOUS* (*AG*) chromatin and to physically interact with the CLF protein (Saleh *et al.*, 2007). We could speculate some interaction between SDG7 and SDG24 at specific flowering loci, as the double mutant *sdg7-5, sdg24-2* displayed an early flowering phenotype under long and short day conditions.

The striking formation of aerial rosettes in the *sdg7-5* mutant is also interesting, as this phenotype is also displayed in delayed flowering mutants of the genes *SOC1*, *AGAMOUS-LIKE 42* (*AGL42*), *AGL71* and *AGL72* (Dorca-Fornell *et al.*, 2011). These genes are implicated in promoting flowering at the shoot apical and axillary meristems. In addition, mutation in *AERIAL ROSETTE-1* (*ART1*) also shows the appearance of aerial rosettes, and its implication in flowering development have been associated to mediating activation of *FLC* expression independently of *FRIGIDA* (*FRI*). The association of SGD7 with the induction of *VIN3* (Lee *et al.*, 2015) might indicate certain implications in flowering time pathways. Further analyses are needed to reveal a possible association of SDG7 and SDG24 in flowering time.

In addition, results from the combinatorial peptide-binding assays suggested that both SDG7 and SDG24 could recognize different PTMs. The presence of acetylated histones H3 and H4 (H3ac and H4ac) and the methylation of the H3K4me1/me2/me3 is characteristic of transcriptionally active chromatin (Sanchez and Gutierrez, 2009; Karnani *et al.*, 2010; Costas *et al.*, 2011). Also, acetylation of H2BK5 has been associated with active genes and the presence of H2AK5ac with double strand breaks (DSBs) (Ikura *et al.*, 2007; Wang *et al.*, 2008). On the other hand, binding to different H3K27 methylated peptides seem antagonistic between SDG7 and SDG24 in terms of chromatin state distribution, as H3K27me1 is enriched in constitutive heterochromatin, whereas H3K27me3 marks gene repression in the euchromatin (Sequeira-Mendes *et al.*, 2014; Sequeira-Mendes and Gutierrez, 2016).

More intriguing was the preferential recognition of H4K20me1/me2/me3 substrates by SDG24. In mammalian cells studies have shown that H4K20 methylation may affect higher-order chromatin structure due to its position at the junction of DNA and the nucleosome (Lu *et al.*, 2008). Furthermore, acetylation of a neighboring residue, H4K16, has been demonstrated to regulate higher-order chromatin dynamics and reduction for H4K20me2 affinity by 53BP1, a target of ATM kinase that forms nuclear foci upon DSB induction (Daley and Sung, 2014). Interestingly, our histone peptide binding of SDG24 showed that in the presence of H4K16ac, the affinity for H4K20me2 decreased by 70%, and 20% for the H4K20me3 mark. However the demonstration of H4K20 in plant chromatin remains elusive. Future advances depend on addressing PTMs changes, considering the particular nucleolar localization of SDG24 and the major alterations of the rDNA repeats in SDG24 mutant.

5.4 SDG7 and SDG24 have a function in the control of rDNA stability

In Arabidopsis actively transcribed NORs associate to the nucleolus differently across accessions, being in some cases the NOR4 or NOR2 mutually exclusive, and in other cases, a combination of both at the same time (Rabanal *et al.*, 2017). In Arabidopsis the study of 45S 3′ETS regions allows the identification of rDNA variants associated specifically with each NOR (Pontvianne *et al.*, 2012; Rabanal *et al.*, 2017). Our work provided evidence that the disruption of *SDG7* and *SDG24* expression affected directly rDNA stability in differentiated organs during the Arabidopsis sporophytic life cycle.

In particular, the SDG24 null mutant, *sdg24-2*, caused the loss of inactive copies accompanied by a loss of telomeric sequences. These results have led us to reason that a major part of NOR2 was lost in the *sdg24-2* allele. Studies in CAF1 mutants (*fas1*, *fas2*) have shown that only ~10–20% of rDNA is sufficient for cell viability in Arabidopsis (Mozgova *et al.*, 2010; Pavlistova *et al.*, 2016), which explained why loosing ~70% of rDNA copies in *sdg24-2* does not represent a major challenge for the plant growth under normal conditions. However, it must kept in mind that cell cycle defects are observed within the RAM, involving a lower cell division capacity and early accumulation of S-phase cells in the proliferative domain of the root. Interestingly, the unbalance in rDNA copies was coped in the *sdg24-2* mutant by relocating all remaining variants inside the nucleolus in a highly decondensed state. By the contrary, *sdg7-5* mutants seemed to alter rDNA stability in a different manner, that is, by an apparent increase of active rDNA copies, although the rRNA silencing mechanism appears to be very similar to the wild type. The **figure 5.2** displays a schematic representation of the results obtained from *sdg7-5* and *sdg24-2* mutants regarding the rDNA and rRNA patterns.

Different mechanism could explain the loss of DNA occurring during the cell cycle. For instance, in mammalian cells reports indicate that breaks in the 45S repeat are predominantly repaired through NHEJ (Harding *et al.*, 2015); although, HR-mediated DNA repair is also involved in the restoration of breaks in the rDNA (van Sluis and McStay, 2015). Particularly, in human cells, the wide distribution of rDNA repeats over multiple chromosomes complicates HR-mediated repair, resulting in repeat loss (Warmerdam *et al.*, 2016). In Arabidopsis, studies in *fas* mutants indicate that the loss of rDNA repeats depends partly on the activity of the DNA double-strand break repair factor RAD51B, a paralog of RAD51, probably mediated via the single-strand annealing (SSA) recombination pathway (Muchova *et al.*, 2015).

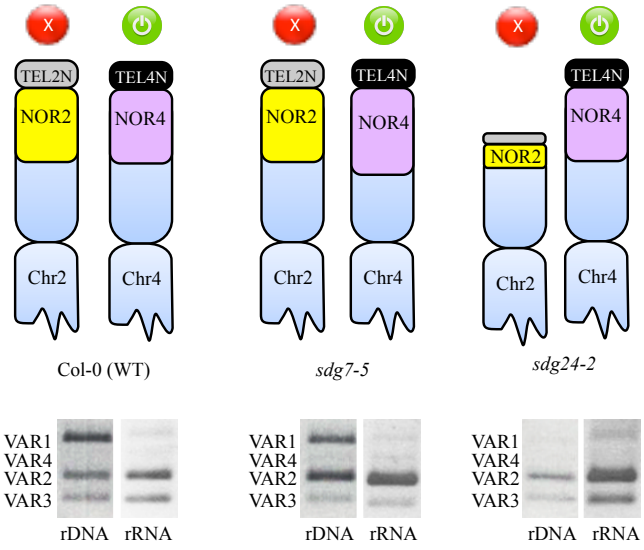


Fig. 5.2. rDNA and rRNA patterns of SDG7 and SDG24 mutants. Representation of rDNA copy number variation of the wild type (Col-0), *sdg7-5* and *sdg24-2* mutants. VAR1 and a subset of VAR3 are located in the NOR2, while VAR2 and the rest of VAR3 in the NOR4. A silencing mechanism inactivates NOR2 variants expression during vegetative development. Red and green circles represent repression or activation of rRNA, respectively. The lower panels show PCR amplification products of rDNA and rRNA templates. Denote the loss of VAR1 copies in *sdg24-2,* and the increase of VAR2 copies in *sdg7-5*. Chr2: Chromosome 2; Chr4: chromosome 4; TEL2N: telomeric sequences of Chr2 short arm; TEL4N: telomeric sequences of Chr2 short arm.

Interestingly, from our expression analysis of components of the DSB pathway, we observed a slight increased of *RAD51* and *BRCA1* expression, and this might correspond to a very localized response of DBS repair at the nucleolar periphery or the basal activity of this proteins.

Another possibility accounting for changes of rDNA copies will be the result of DSBs in the replication fork. In the budding yeast, breaks within rDNA due to the activity of FOB1 can be repaired by equal or unequal sister chromatid recombination (ESCR or USCR, respectively), being USCR the pathway leading to rDNA copy number changes (**Fig. 5.3**) (Kobayashi *et al.*, 1998; Ganley *et al.*, 2009).

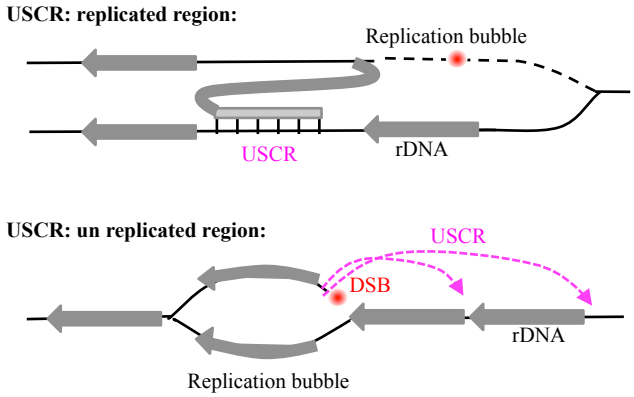


Figure 5.3. Copy number changes causes by USCR. A DSB (red dot) forms in the rDNA replication bubble and recombination could follow different pathways. USCR in replicated regions leads to increase of copy number while USCR recombination with the unreplicated (nonbubble) region results in loss of rDNA copy number (pink lines) and formation of rDNA circles. Grey arrows represent each unit of rDNA. Adapted from Ganley et al., 2009.

When a broken end is recombined within the replicated region in an USCR fashion, duplication of genes can occur, increasing therefore copy number. On the other hand, if a broken end recombined with its own chromatid in the unreplicated region (intrachromatid recombination), an extrachromosomal rDNA circle (ERC) is produced, resulting in a "looping out" of rDNA copies from the chromatid (Ganley *et al.*, 2009). Extra chromosomal DNA circles have been found in many organisms, including plants, although we have not determined yet if this is the case in the *sdg7-5* and *sdg24-2* mutants. In Arabidopsis, the size of circular DNA ranges between 2 kb and 20 kb, which is similar to the sizes found in other organisms (Cohen *et al.*, 2008; Cohen *et al.*, 2010). We can speculate that in the absence of SDG24, looping out of several units of rDNA copies of NOR2 might occur within several rounds of replication during the seed formation.

In Arabidopsis, a recent study published by Mohannath et al. (2016) uncovers the mechanism by which *atxr5/6* mutant alters rDNA patterns of normally silenced variants (NOR2). The overrepresentation of VAR1 in *atxr5/6* mutant responds to a major recombination event, where much of the NOR4 and its associated telomere sequences (TEL4N) have been lost and substituted with the corresponding sequences of chromosome 2. Therefore NOR2-derived VAR1 genes at their new location (Chr4) scape silencing, while NOR2 remained inactivated (**Fig. 5.4**). However this arrangement was independent of the *atxr5/6* mutations, as some *atxr5/6* mutant lines also displayed the wild type rDNA pattern (Mohannath *et al.*, 2016).

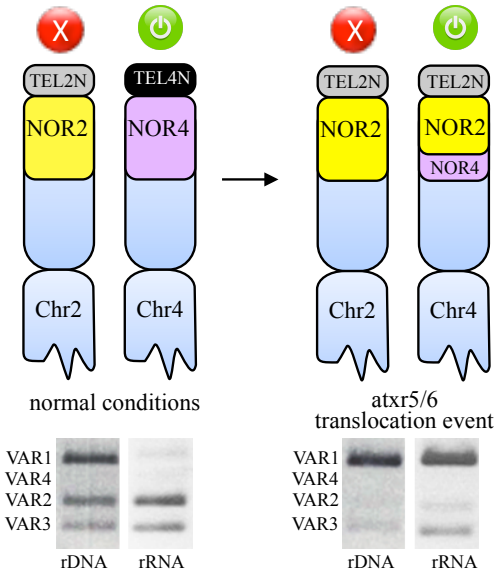


Fig. 5.4. Recombination event explaining *atxr5/6* RDNA patterns. A recombinant event translocate NOR2-derived rDNA genes (VAR1) and its Telomere-associated sequences into the NOR4 location, replacing NOR4 variants (VAR2-VAR3) in the chromosome 4, independently of *atxr5/6* mutation. Red and green circles represent repression and activation of rRNA, respectively. Chr2: Chromosome 2; Chr4: chromosome 4; TEL2N: telomeric sequences of Chr2 short arm; TEL4N: telomeric sequences of Chr2 short arm. Modified from Mohannath et al., 2016.

More than one mechanism could account for replacement of NOR4 variants, including one single crossover between NOR2 and NOR4 during meiosis, or by a non-reciprocal translocation resulted from a break-induced replication initiated by a collapsed replication fork or DSBs. If DBS repair within NOR4 is mediated by homologous NOR2 sequences rather than a NOR4 unbroken sister chromatid, the result from this recombination event will not alter the amount of rDNA copies, which is in agreement with ongoing experiments on *atxr5/6* mutant derived from our work (data not shown).

Form our work, we also observed that the rDNA pattern given by the *sdg24-2* mutant was dominant over the *sdg7-5* mutation, and also over *atxr5/6*. However, in the context of nucleolar dominance, where NOR-derived rDNAs from one progenitor are selectively silenced, *sdg24-2* falls into another category, since the loss of NOR2-rDNA copies seems to be dominant over the silencing mechanisms upon NOR2. This phenomenon uncovers an earlier event controlling the stability of rDNA copies. Overall, although the mechanism of NOR2 silencing is still present in the *atxr5/6* mutant, the phenotype of the *atxr5/6, sdg7-5* triple mutant also responds to a translocation event of NOR2-derived variants into the NOR4, as the triple mutant displayed overrepresentation of VAR1 rDNA copies. On the contrary, in the triple mutant *atxr5/6, sdg24-2* various possible outcomes could result from combining different NOR2 and NOR4 alleles. However it became evident that the *sdg24-2* mutation exerts control over the wild type NOR2 and the translocated NOR2 into the NOR4 location present in the *atxr5/6* mutant.

It has been proposed that transcriptional regulation of NOR2 is not the result of silencing of individual rRNA genes based on their sequences, but rather the consequence of a single chromosomal event (Chandrasekhara *et al.*, 2016; Mohannath *et al.*, 2016). Differences in both NOR-adjacent sequences have pointed out in the NOR2 to a region composed of transposable elements and transposon remnants that spans for ~60 kb before the first protein-coding sequence, while in the NOR4 only a 3 kb region separates the last rDNA gene from the next protein-coding sequence.

Taking together our results, we propose that both SDG7 and SDG24 might regulate the proper DNA damage repair mechanisms exclusively associated to rDNA regions. Whether this mechanism is mediated by the deposition of a specific mark in the rDNA chromatin is yet to be uncovered.

Conclusions
Conclusiones

CONCLUSIONS

1- SDG7 is as nuclear and cytoplasmic protein, expressed preferentially in proliferative cells of Arabidopsis organs.

2- The *SDG24* gene possesses several mRNA splicing variants expressed in different organs. The *SDG24.1* variant is expressed in the primary and lateral roots, trichomes, anthers and gametic cells of mature ovules. Expression of *SDG24.2* is found only in the anthers.

3- The SDG24 protein (translated from *SDG24.1* variant) shows nucleolar localization along the root meristem cells.

4- Both SDG7 and SDG24 are absent during early events of mitosis, and loaded into the nucleus during telophase in root meristem cells.

5- SDG7 binds preferentially to the modified histones residues H3K4ac, H3K9me2 and H3K27me1, in a context where H3R8/R17/R26 are dimethylated and histones H2A, H2B and H4 are acetylated. SDG24 shows a high binding capacity for acetylated H2A and H2B, H3K27me2/me3, and H4K20me1/me2, which is repressed by nearby H4K16ac.

6- Both SDG7 and SDG24 are SET domain-containing proteins with no detectable histone methyltransferase activity using *in vitro* assays with proteins produced in *E. coli*.

7- The absence of SDG7 and SDG24 in the null alleles *sdg7-5* and *sdg24-2* produces alterations in the cell division rate, the DNA replication pattern and the root meristem length.

8- The rDNA pattern is severely altered in the *sdg7-5* and *sdg24-2* mutants. The *sdg7-5* mutant increases NOR4-active rDNA VAR2, whereas the *sdg24-2* mutant displays a loss of NOR2-inactive rDNA VAR1 and a shortening of telomeres. Both, *sdg7-5* and *sdg24-2* mutants do not alter rRNA silencing patterns in mature organs.

9- The loss of NOR2-inactive rDNA variants in the *sdg24-2* mutant is epistatic to *sdg7-5* and *atxr5/6* mutations regarding the rDNA phenotypes.

CONCLUSIONES

1- La proteína SDG7 posee una localización nuclear y citoplasmática, expresada preferencialmente en las células proliferativas de los órganos de Arabidopsis.

2- El gen de *SDG24* posee diferentes variantes de splicing expresadas en diferentes órganos. La variante *SDG24.1* se expresa en la raíz principal y laterales, los tricomas, las anteras, y en las células gaméticas de óvulos maduros. La expresión de *SDG24.2* se localiza solo en las anteras.

3- La proteína SDG24 (traducida de la variante *SDG24.1*) muestra una localización nucleolar en células del meristemo de la raíz.

4- Las proteínas SDG7 y SDG24 se encuentran ausentes desde el principio de la mitosis, y vuelven a ser cargadas al núcleo durante la telofase en células del meristemo de la raíz.

5- SDG7 se une preferentemente a los residuos de histonas modificados H3K4ac, H3K9me2 y H3K27me1, en un contexto en el que existen dimetilaciones de H3R8/R17/R26 y acetilación de las histonas H2A, H2B y la H4. SDG24 presentó una afinidad de unión por residuos H2A y H2B acetilados, H3K27me2/me3, y H4K20me1/me2, unión que se ve disminuida por la presencia de H4K16ac.

6- Mediante ensayos *in vitro* con proteínas SDG7 y SDG24 purificadas de *E. coli* no se detectó actividad metiltransferasa sobre histonas.

7- La ausencia de expresión de SDG7 y SDG24 en los mutantes nulos *sdg7-5 y sdg24-2* produce alteraciones del patrón de división celular, de la replicación del DNA y de la longitud del meristemo de la raíz.

8- El patrón de rDNA se encuentra severamente afectado en los mutantes *sdg7-5 y sdg24-2*. El mutante *sdg7-5* presenta un incremento de la variante activa de rDNA (VAR2) asociada al NOR4, mientras que el mutante *sdg24-2* exhibe una pérdida de la variante inactiva de rDNA (VAR1) asociada al NOR4 y un acortamiento de los telómeros. Ambos mutantes no alteran el patrón de silenciamiento del rRNA observado en órganos maduros.

9- La pérdida de variantes de rDNA inactivas del NOR2 del mutante *sdg24-2* es epistática *sdg7-5* y *atxr5/6* respecto a sus patrones de rDNA.

References

Abou-Ellail, M., Cooke, R., and Saez-Vasquez, J. (2011). Variations in a team: major and minor variants of Arabidopsis thaliana rDNA genes. Nucleus *2*, 294-299.

Agusti, J., and Greb, T. (2013). Going with the wind--adaptive dynamics of plant secondary meristems. Mech Dev *130*, 34-44.

Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., Nussaume, L., Noh, Y.S., Amasino, R., and Scheres, B. (2004). The PLETHORA genes mediate patterning of the Arabidopsis root stem cell niche. Cell *119*, 109-120.

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., and Raff, M. (2014). Essential Cell Biology, 4th edn (New York: Garland-Science). pp171-176.

Alvarez-Buylla, E.R., Benitez, M., Corvera-Poire, A., Chaos Cador, A., de Folter, S., Gamboa de Buen, A., Garay-Arroyo, A., Garcia-Ponce, B., Jaimes-Miranda, F., Perez-Ruiz, R.V.*, et al.* (2010). Flower development. Arabidopsis Book *8*, e0127.

Balcerowicz, D., Schoenaers, S., and Vissenberg, K. (2015). Cell fate determination and the switch from diffuse growth to planar polarity in Arabidopsis root epidermal cells. Front Plant Sci *6*, 1-13.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Baroux, C., and Grossniklaus, U. (2015). The maternal-to-zygotic transition in flowering plants: evidence, mechanisms, and plasticity. Curr Top Dev Biol *113*, 351-371.

Baroux, C., Raissig, M.T., and Grossniklaus, U. (2011). Epigenetic regulation and reprogramming during gamete formation in plants. Curr Opin Genet Dev *21*, 124-133.

Baumbusch, L., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G., and Aalen, R.B. (2001). The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res *29*, 4319-4333.

Belmonte, M.F., Kirkbride, R.C., Stone, S.L., Pelletier, J.M., Bui, A.Q., Yeung, Hashimoto, M., Fei, J.F., Harada, C.M., Munoz, M.D.*, et al.* (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. Proc Natl Acad Sci U S A *110*, E435-E444.

Bemer, M., and Grossniklaus, U. (2012). Dynamic regulation of Polycomb group activity during plant development. Curr Opin Plant Biol *15*, 523-529.

Benfey, P., and Schiefelbein, J.W. (1994). Insights into root development from Arabidopsis root mutants. Plant Cell Environ *17*, 675-680.

Benoit, M., Layat, E., Tourmente, S., and Probst, A.V. (2013). Heterochromatin dynamics during developmental transitions in Arabidopsis - a focus on ribosomal DNA loci. Gene *526*, 39-45.

Berger, F. (1999). Endosperm development. Curr Opin Plant Biol *2*, 28-32.

Berger, F., Hung, C.Y., Dolan, L., and Schiefelbein, J. (1998). Control of cell division in the root epidermis of Arabidopsis thaliana. Dev Biol *194*, 235-245.

Berger, F., and Twell, D. (2011). Germline specification and function in plants. Annu Rev Plant Biol *62*, 461-484.

Berr, A., McCallum, E.J., Menard, R., Meyer, D., Fuchs, J., Dong, A., and Shen, W.H. (2010). Arabidopsis SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. Plant Cell *22*, 3232-3248.

Berr, A., Shafiq, S., Pinon, V., Dong, A., and Shen, W.H. (2015). The trxG family histone methyltransferase SET DOMAIN GROUP 26 promotes flowering via a distinctive genetic pathway. Plant J *81*, 316-328.

Berr, A., Xu, L., Gao, J., Cognat, V., Steinmetz, A., Dong, A., and Shen, W.H. (2009). SET DOMAIN GROUP25 encodes a histone methyltransferase and is involved in FLOWERING LOCUS C activation and repression of flowering. Plant Physiol *151*, 1476-1485.

Bonenfant, D., Coulot, M., Towbin, H., Schindler, P., and van Oostrum, J. (2006). Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. Mol Cell Proteomics *5*, 541-552.

Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.M., and Lamond, A.I. (2010). The nucleolus under stress. Mol Cell *40*, 216-227.

Bowman, G.D., and Poirier, M.G. (2015). Post-translational modifications of histones that influence nucleosome dynamics. Chem Rev *115*, 2274-2295.

Bowman, J.L. (1994). Flower development. In Arabidopsis: an atlas of morphology and development (New York: Springer-Verlag), pp. 276-278.

Calarco, J.P., Borges, F., Donoghue, M.T., Van Ex, F., Jullien, P.E., Lopes, T., Gardner, R., Berger, F., Feijo, J.A., Becker, J.D*., et al.* (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell *151*, 194-205.

Capron, A., Chatfield, S., Provart, N., and Berleth, T. (2009). Embryogenesis: pattern formation from a single cell. Arabidopsis Book *7*, e0126.

Caro, E., Stroud, H., Greenberg, M.V., Bernatavichute, Y.V., Feng, S., Groth, M., Vashisht, A.A., Wohlschlegel, J., and Jacobsen, S.E. (2012). The SET-domain protein SUVR5 mediates H3K9me2 deposition and silencing at stimulus response genes in a DNA methylation-independent manner. PLoS Genet *8*, e1002995.

Casamitjana-Martínez, E., Hofhuis, H.F., Xu, J., Liu, C.-M., Heidstra, R., and Scheres, B. (2003). Root-specific CLE19 overexpression and the sol1/2 suppressors implicate a CLV-like pathway in the control of Arabidopsis root meristem maintenance. Current Biol *13*, 1435-1441.

Casas-Mollano, J.A., Lao, N.T., and Kavanagh, T.A. (2006). Intron-regulated expression of SUVH3, an Arabidopsis Su(var)3-9 homologue. J Exp Bot *57*, 3301-3311.

Causier, B., Ashworth, M., Guo, W., and Davies, B. (2012). The TOPLESS interactome: a framework for gene repression in Arabidopsis. Plant Physiol *158*, 423-438.

Cazzonelli, C.I., Nisar, N., Roberts, A.C., Murray, K.D., Borevitz, J.O., and Pogson, B.J. (2014). A chromatin modifying enzyme, SDG8, is involved in morphological, gene expression, and epigenetic responses to mechanical stimulation. Front Plant Sci *5*, 533.

Chandrasekhara, C., Mohannath, G., Blevins, T., Pontvianne, F., and Pikaard, C.S. (2016). Chromosome-specific NOR inactivation explains selective rRNA gene silencing and dosage control in Arabidopsis. Genes Dev *30*, 177-190.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. Development *131*, 5263-5276.

Chen, L., Luo, J., Cui, Z., Xue, M., Wang, L., Zhang, X., Pawlowski, W.P., and He, Y. (2017). ATX3, ATX4, and ATX5 encode putative H3K4 methyltransferases and are critical for plant development. Plant Physiol *174*, 1795-1806.

Chen, Z.J., and Pikaard, C.S. (1997). Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance. Genes Dev *11*, 2124-2136.

Cho, L.H., Yoon, J., and An, G. (2017). The control of flowering time by environmental factors. Plant J *90*, 708-719.

Cloix, C., Tutois, S., Mathieu, O., Cuvillier, C., Espagnol, M.C., Picard, G., and Tourmente, S. (2000). Analysis of 5S rDNA arrays in Arabidopsis thaliana: physical mapping and chromosome-specific polymorphisms. Genome Res *10*, 679-690.

Cohen, S., Agmon, N., Sobol, O., and Segal, D. (2010). Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. Mob DNA *1*, 1-11.

Cohen, S., Houben, A., and Segal, D. (2008). Extrachromosomal circular DNA derived from tandemly repeated genomic sequences in plants. Plant J *53*, 1027-1034.

Colon-Carmona, A., You, R., Haimovitch-Gal, T., and Doerner, P. (1999). Technical advance: spatio-temporal analysis of mitotic activity with a labile cyclin-GUS fusion protein. Plant J *20*, 503-508.

Congdon, L.M., Houston, S.I., Veerappan, C.S., Spektor, T.M., and Rice, J.C. (2010). PR-Set7-mediated monomethylation of histone H4 lysine 20 at specific genomic regions induces transcriptional repression. J Cell Biochem *110*, 609-619.

Copenhaver, G.P., and Pikaard, C. (1996). Two-dimentional RFLP analyses reveal megabase-size clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. Plant J *9*, 273-282.

Costas, C., Sanchez, M.d.L., Stroud, H., Yu, Y., Oliveros, J.C., Feng, S., Benguria, A., Lopez-Vidriero, I., Zhang, X., Solano, R.*, et al.* (2011). Genome-wide mapping of Arabidopsis thaliana origins of DNA replication and their associated epigenetic marks. Nat Struct Mol Biol *18*, 395-400.

Czesnick, H., and Lenhard, M. (2015). Size control in plants--lessons from leaves and flowers. Cold Spring Harb Perspect Biol *7*, a019190.

Daley, J.M., and Sung, P. (2014). 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. Mol Cell Biol *34*, 1380-1388.

Deal, R.B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. Dev Cell *18*, 1030-1040.

Denay, G., Chahtane, H., Tichtinsky, G., and Parcy, F. (2017). A flower is born: an update on Arabidopsis floral meristem formation. Curr Opin Plant Biol *35*, 15-22.

Desvoyes, B., Fernandez-Marcos, M., Sequeira-Mendes, J., Otero, S., Vergara, Z., and Gutierrez, C. (2014). Looking at plant cell cycle from the chromatin window. Front Plant Sci *5*, 369.

Desvoyes, B., Ramirez-Parra, E., Xie, Q., Chua, N.H., and Gutierrez, C. (2006). Cell type-specific role of the retinoblastoma/E2F pathway during Arabidopsis leaf development. Plant Physiol *140*, 67-80.

Ding, Y., Avramova, Z., and Fromm, M. (2011a). The Arabidopsis trithorax-like factor ATX1 functions in dehydration stress responses via ABA-dependent and ABA-independent pathways. Plant J *66*, 735-744.

Ding, Y., Avramova, Z., and Fromm, M. (2011b). Two distinct roles of ARABIDOPSIS HOMOLOG OF TRITHORAX1 (ATX1) at promoters and within transcribed regions of ATX1-regulated genes. Plant Cell *23*, 350-363.

Dong, G., Ma, D.P., and Li, J. (2008). The histone methyltransferase SDG8 regulates shoot branching in Arabidopsis. Biochem Biophys Res Commun *373*, 659-664.

Dong, H., Liu, D., Han, T., Zhao, Y., Sun, J., Lin, S., Cao, J., Chen, Z.H., and Huang, L. (2015). Diversification and evolution of the SDG gene family in Brassica rapa after the whole genome triplication. Sci Rep *5*, 16851.

Dorca-Fornell, C., Gregis, V., Grandi, V., Coupland, G., Colombo, L., and Kater, M.M. (2011). The Arabidopsis SOC1-like genes AGL42, AGL71 and AGL72 promote flowering in the shoot apical and axillary meristems. Plant J *67*, 1006-1017.

Doyle, J.J., and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. Focus *12*, 13-15.

Draizen, E.J., Shaytan, A.K., Marino-Ramirez, L., Talbert, P.B., Landsman, D., and Panchenko, A.R. (2016). HistoneDB 2.0: a histone database with variants--an integrated resource to explore histones and their variants. Database (Oxford) *2016*, 1-10.

Dutt, M., Dhekney, S.A., Soriano, L., Kandel, R., and Grosser, J.W. (2014). Temporal and spatial control of gene expression in horticultural crops. Hortic Res *1*, 14047.

Dvorackova, M., Fojtova, M., and Fajkus, J. (2015). Chromatin dynamics of plant telomeres and ribosomal genes. Plant J *83*, 18-37.

Dvorackova, M., Raposo, B., Matula, P., Fuchs, J., Schubert, V., Peska, V., Desvoyes, B., Gutierrez, C., and Fajkus, J. (2017). Replication of ribosomal DNA in Arabidopsis occurs both inside and outside the nucleolus during S phase progression. J Cell Sci, DOI: 10.1242/jcs.202416.

Earley, K.W., Pontvianne, F., Wierzbicki, A.T., Blevins, T., Tucker, S., Costa-Nunes, P., Pontes, O., and Pikaard, C.S. (2010). Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. Genes Dev *24*, 1119-1132.

Ebbs, M.L., and Bender, J. (2006). Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell *18*, 1166-1176.

Edgar, B.A., Zielke, N., and Gutierrez, C. (2014). Endocycles: a recurrent evolutionary innovation for post-mitotic cell growth. Nat Rev Mol Cell Biol *15*, 197-210.

Feng, W., Hale, C.J., Over, R.S., Cokus, S.J., Jacobsen, S.E., and Michaels, S.D. (2017). Large-scale heterochromatin remodeling linked to overreplication-associated DNA damage. Proc Natl Acad Sci U S A *114*, 406-411.

Fernandez-Marcos, M., Desvoyes, B., Manzano, C., Liberman, L.M., Benfey, P.N., Del Pozo, J.C., and Gutierrez, C. (2017). Control of Arabidopsis lateral root primordium boundaries by MYB36. New Phytol *213*, 105-112.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.K., and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome Res *20*, 45-58.

Fischle, W., Wang, Y., and Allis, C.D. (2003). Histone and chromatin cross-talk. Curr Opin Cell Biol *15*, 172-183.

Fletcher, J.C. (2002). Shoot and floral meristem maintenance in arabidopsis. Annu Rev Plant Biol *53*, 45-66.

Forzani, C., Aichinger, E., Sornay, E., Willemsen, V., Laux, T., Dewitte, W., and Murray, J.A. (2014). WOX5 suppresses CYCLIN D activity to establish quiescence at the center of the root stem cell niche. Curr Biol *24*, 1939-1944.

Galbraith, D.W. (2009). Simultaneous flow cytometric quantification of plant nuclear DNA contents over the full range of described angiosperm 2C values. Cytometry A *75*, 692-698.

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., Scheres, B., root, P.p.a.d.-d.m.r.o.A., and development. Nature 449, -. (2007). PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. Nature, 1045-1057.

Ganley, A.R., Ide, S., Saka, K., and Kobayashi, T. (2009). The effect of replication initiation on gene amplification in the rDNA and its relationship to aging. Mol Cell *35*, 683-693.

Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science *324*, 1447-1451.

Geldner, N., and Salt, D.E. (2014). Focus on roots. Plant Physiol *166*, 453-454.

Gonzalez, N., Vanhaeren, H., and Inze, D. (2012). Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci *17*, 332-340.

Grini, P.E., Thorstensen, T., Alm, V., Vizcay-Barrena, G., Windju, S.S., Jorstad, T.S., Wilson, Z.A., and Aalen, R.B. (2009). The ASH1 HOMOLOG 2 (ASHH2) histone H3 methyltransferase is required for ovule and anther development in Arabidopsis. PLoS One *4*, e7817.

Guo, L., Yu, Y., Law, J.A., and Zhang, X. (2010). SET DOMAIN GROUP2 is the major histone H3 lysine 4 trimethyltransferase in Arabidopsis. Proc Natl Acad Sci U S A *107*, 18557-18562.

Gurard-Levin, Z.A., and Almouzni, G. (2014). Histone modifications and a choice of variant: a language that helps the genome express itself. F1000Prime Rep *6*, 1-10.

Gutierrez, C. (2009). The Arabidopsis Cell Division Cycle. Arabidopsis Book *7*, e0120.

Haga, N., Kato, K., Murase, M., Araki, S., Kubo, M., Demura, T., and al., e. (2007). R1R2R3-Myb proteins positively regulate cytokinesis through activation of KNOLLE transcription in Arabidopsis thaliana. Development *134*, 1101-1110.

Haga, N., Kobayashi, K., Suzuki, T., Maeo, K., Kubo, M., Ohtani, M., Mitsuda, N., Demura, T., Nakamura, K., Jurgens, G.*, et al.* (2011). Mutations in MYB3R1 and MYB3R4 cause pleiotropic developmental defects and preferential down-regulation of multiple G2/M-specific genes in Arabidopsis. Plant Physiol *157*, 706-717.

Han, Y.F., Dou, K., Ma, Z.Y., Zhang, S.W., Huang, H.W., Li, L., Cai, T., Chen, S., Zhu, J.K., and He, X.J. (2014). SUVR2 is involved in transcriptional gene silencing by associating with SNF2-related chromatin-remodeling proteins in Arabidopsis. Cell Res *24*, 1445-1465.

Harding, S.M., Boiarsky, J.A., and Greenberg, R.A. (2015). ATM Dependent Silencing Links Nucleolar Chromatin Reorganization to DNA Damage Recognition. Cell Rep *13*, 251-259.

Havlova, K., Dvorackova, M., Peiro, R., Abia, D., Mozgova, I., Vansacova, L., Gutierrez, C., and Fajkus, J. (2016). Variation of 45S rDNA intergenic spacers in Arabidopsis thaliana. Plant Mol Biol *92*, 457-471.

Hayashi, K., Hasegawa, J., and Matsunaga, S. (2013). The boundary of the meristematic and elongation zones in roots: endoreduplication precedes rapid cell expansion. Sci Rep *3*, DOI: 10.1038/srep02723.

He, G., Elling, A.A., and Deng, X.W. (2011). The epigenome and plant development. Annu Rev Plant Biol *62*, 411-435.

Hennig, L., and Derkacheva, M. (2009). Diversity of Polycomb group complexes in plants: same rules, different players? Trends Genet *25*, 414-423.

Hernandez-Garcia, C.M., and Finer, J.J. (2014). Identification and validation of promoters and cis-acting regulatory elements. Plant Sci *217-218*, 109-119.

Hernandez-Verdun, D. (2011a). Assembly and disassembly of the nucleolus during the cell cycle. Nucleus *2*, 189-194.

Hernandez-Verdun, D. (2011b). Structural Organization of the nucleolus as a consequence of the dynamics of ribosome biogenesis. In The Nucleolus, M. Olson, ed. (New York: Springer), pp. 1-28.

Herz, H.M., Garruss, A., and Shilatifard, A. (2013). SET for life: biochemical activities and biological functions of SET domain-containing proteins. Trends Biochem Sci *38*, 621-639.

Heyman, J., Cools, T., Vandenbussche, F., Heyndrickx, K.S., Van Leene, J., Vercauteren, I., Vanderauwera, S., Vandepoele, K., De Jaeger, G., Van Der Straeten, D.*, et al.* (2013). ERF115 controls root quiescent center cell division and stem cell replenishment. Science *342*, 860-863.

Hoppmann, V., Thorstensen, T., Kristiansen, P.E., Veiseth, S.V., Rahman, M.A., Finne, K., Aalen, R.B., and Aasland, R. (2011). The CW domain, a new histone recognition module in chromatin proteins. EMBO J *30*, 1939-1952.

Hsieh, J., and Gage, F.H. (2005). Chromatin remodeling in neural development and plasticity. Curr Opin Cell Biol *17*, 664-671.

Ikura, T., Tashiro, S., Kakino, A., Shima, H., Jacob, N., Amunugama, R., Yoder, K., Izumi, S., Kuraoka, I., Tanaka, K.*, et al.* (2007). DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics. Mol Cell Biol *27*, 7028-7040.

Ito, M., Iwase, M., Kodama, H., Lavisse, P., Komamine, A., Nishihama, R., and al., e. (1998). A novel cis-acting element in promoters of plant B-type cyclin genes activates M phase-specific transcription. Plant Cell *10*, 331-341.

Jackson, J., Lindroth, M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556-560.

Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S.E. (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma *122*, 308-315.

Jacob, Y., Feng, S., LeBlanc, C.A., Bernatavichute, Y.V., Stroud, H., Cokus, S., Johnson, L.M., Pellegrini, M., Jacobsen, S.E., and Michaels, S.D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat Struct Mol Biol *16*, 763-768.

Jeong, C.W., Roh, H., Dang, T.V., Choi, Y.D., Fischer, R.L., Lee, J.S., and Choi, Y. (2011). An E3 ligase complex regulates SET-domain polycomb group protein activity in Arabidopsis thaliana. Proc Natl Acad Sci U S A *108*, 8036-8041.

Jeong, S., and Clark, S.E. (2005). Photoperiod regulates flower meristem development in Arabidopsis thaliana. Genetics *169*, 907-915.

Jing, Y., Sun, H., Yuan, W., Wang, Y., Li, Q., Liu, Y., Li, Y., and Qian, W. (2016). SUVH2 and SUVH9 couple two essential steps for transcriptional gene silencing in arabidopsis. Mol Plant *9*, 1156-1167.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. (2007). The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr Biol *17*, 379-384.

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. (2008). SRA-domain proteins required for DRM2-mediated de novo DNA methylation. PLoS Genet *4*, e1000280.

Jorgensen, S., Schotta, G., and Sorensen, C.S. (2013). Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. Nucleic Acids Res *41*, 2797-2806.

Jullien, P.E., and Berger, F. (2010). DNA methylation reprogramming during plant sexual reproduction? Trends Genet *26*, 394-399.

Jung, J.K., and McCouch, S. (2013). Getting to the roots of it: Genetic and hormonal control of root architecture. Front Plant Sci *4*, 186.

Karnani, N., Taylor, C.M., Malhotra, A., and Dutta, A. (2010). Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. Mol Biol Cell *21*, 393-404.

Kawashima, T., and Berger, F. (2014). Epigenetic reprogramming in plant sexual reproduction. Nat Rev Genet *15*, 613-624.

Ko, J.H., Mitina, I., Tamada, Y., Hyun, Y., Choi, Y., Amasino, R.M., Noh, B., and Noh, Y.S. (2010). Growth habit determination by the balance of histone methylation activities in Arabidopsis. EMBO J *29*, 3208-3215.

Kobayashi, T., Heck, D.J., Nomura, M., and Horiuchi, T. (1998). Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. Genes Dev *12*, 3821-3830.

Kohler, C., and Kradolfer, D. (2011). Epigenetic mechanisms in the endosperm and their consequences for the evolution of flowering plants. Biochim Biophys Acta *1809*, 438-443.

Kohler, C., Page, D.R., Gagliardini, V., and Grossniklaus, U. (2005). The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. Nature Genetics *37*, 28-30.

Kohler, C., and Springer, N. (2017). Plant epigenomics-deciphering the mechanisms of epigenetic inheritance and plasticity in plants. Genome Biol *18*, 132.

Koornneef, M., and Meinke, D. (2010). The development of Arabidopsis as a model plant. Plant J *61*, 909-921.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Krauss, V., and Reuter, G. (2000). Two genes become one: the genes encoding heterochromatin protein SU(VAR)3-9 and translation initiation factor subunit eIF-2g are joined to a dicistronic unit in holometabolic insects. Genetics *156*, 1157-1167.

Krichevsky, A., Gutgarts, H., Kozlovsky, S.V., Tzfira, T., Sutton, A., Sternglanz, R., Mandel, G., and Citovsky, V. (2007). C2H2 zinc finger-SET histone methyltransferase is a plant-specific chromatin modifier. Dev Biol *303*, 259-269.

Kuhlmann, M., and Mette, M.F. (2012). Developmentally non-redundant SET domain proteins SUVH2 and SUVH9 are required for transcriptional gene silencing in Arabidopsis thaliana. Plant Mol Biol *79*, 623-633.

Kumpf, R., Thorstensen, T., Rahman, M.A., Heyman, J., Nenseth, H.Z., Lammens, T., Herrmann, U., Swarup, R., Veiseth, S.V., Emberland, G.*, et al.* (2014). The ASH1-RELATED3 SET-domain protein controls cell division competence of the meristem and the quiescent center of the Arabidopsis primary root. Plant Physiol *166*, 632-643.

Larkin, J.C., Oppenheimer, D.G., Pollock, S., and Marks, D. (1993). Arabidopsis GLABROUS7 gene requires downstream sequences for function. the Plant Cell *5*, 1739-1748.

Lavenus, J., Goh, T., Roberts, I., Guyomarc'h, S., Lucas, M., De Smet, I., Fukaki, H., Beeckman, T., Bennett, M., and Laplaze, L. (2013). Lateral root development in Arabidopsis: fifty shades of auxin. Trends Plant Sci *18*, 450-458.

Lawit, S.J., Chamberlin, M.A., Agge, A., Caswell, E., and Albertsen, M.C. (2013). Transgenic manipulation of plant embryo sacs tracked through cell-type-specific fluorescent markers: cell labeling, cell ablation, and adventitious embryos. Plant Reprod *26*, 125-137.

Lawrence, R.J., Earley, K., Pontes, O., Silva, M., Chen, Z.J., Neves, N., Viegas, W., and Pikaard, C.S. (2004). A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. Mol Cell *13*, 599-609.

Lee, J., Yun, J.Y., Zhao, W., Shen, W.H., and Amasino, R.M. (2015). A methyltransferase required for proper timing of the vernalization response in Arabidopsis. Proc Natl Acad Sci U S A *112*, 2269-2274.

Li, S., Liu, L., Li, S., Gao, L., Zhao, Y., Kim, Y.J., and Chen, X. (2016). SUVH1, a Su(var)3-9 family member, promotes the expression of genes targeted by DNA methylation. Nucleic Acids Res *44*, 608-620.

Liu, B., Berr, A., Chang, C., Liu, C., Shen, W.H., and Ruan, Y. (2016). Interplay of the histone methyltransferases SDG8 and SDG26 in the regulation of transcription and plant flowering and development. Biochim Biophys Acta *1859*, 581-590.

Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. Annu Rev Plant Biol *61*, 395-420.

Liu, Z.W., Shao, C.R., Zhang, C.J., Zhou, J.X., Zhang, S.W., Li, L., Chen, S., Huang, H.W., Cai, T., and He, X.J. (2014). The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. PLoS Genet *10*, e1003948.

Lo, S.J., Lee, C.C., and Lai, H.J. (2006). The nucleolus: reviewing oldies to have new understandings. Cell Res *16*, 530-538.

Lofke, C., Scheuring, D., Dunser, K., Scholler, M., Luschnig, C., and Kleine-Vehn, J. (2015). Tricho- and atrichoblast cell files show distinct PIN2 auxin efflux carrier exploitations and are jointly required for defined auxin-dependent root organ growth. J Exp Bot *66*, 5103-5112.

Lu, X., Simon, M.D., Chodaparambil, J.V., Hansen, J.C., Shokat, K.M., and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. Nat Struct Mol Biol *15*, 1122-1124.

MacAlpine, D.M., and Almouzni, G. (2013). Chromatin and DNA replication. Cold Spring Harb Perspect Biol *5*, a010207.

Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). Chromatin as dynamic 10-nm fibers. Chromosoma *123*, 225-237.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Kohler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. EMBO Rep *7*, 947-952.

Manavella, P.A., Koenig, D., Rubio-Somoza, I., Burbano, H.A., Becker, C., and Weigel, D. (2013). Tissue-specific silencing of Arabidopsis SU(VAR)3-9 HOMOLOG8 by miR171a. Plant Physiol *161*, 805-812.

McClintock, B. (1934). The realtionship of a particular chromosomal element to de develpoment of nucleoli in Zea mayz. Z Zellforsch Mikrosk Anat *21*, 294-328.

McCormick, S. (2004). Control of male gametophyte development. Plant Cell *16 Suppl*, 142-153.

Mekhail, K., Seebacher, J., Gygi, S.P., and Moazed, D. (2008). Role for perinuclear chromosome tethering in maintenance of genome stability. Nature *456*, 667-670.

Menges, M., and Murray, J.A. (2002). Synchronous Arabidopsis suspension cultures for analysis of cell-cycle gene activity. Plant J *30*, 203-212.

Milite, C., Feoli, A., Viviano, M., Rescigno, D., Cianciulli, A., Balzano, A.L., Mai, A., Castellano, S., and Sbardella, G. (2016). The emerging role of lysine methyltransferase SETD8 in human diseases. Clin Epigenetics *8*, DOI: 10.1186/s13148-016-0268-4.

Mohannath, G., Pontvianne, F., and Pikaard, C.S. (2016). Selective nucleolus organizer inactivation in Arabidopsis is a chromosome position-effect phenomenon. Proc Natl Acad Sci U S A *113*, 13426-13431.

Mozgova, I., Mokros, P., and Fajkus, J. (2010). Dysfunction of chromatin assembly factor 1 induces shortening of telomeres and loss of 45S rDNA in Arabidopsis thaliana. Plant Cell *22*, 2768-2780.

Muchova, V., Amiard, S., Mozgova, I., Dvorackova, M., Gallego, M.E., White, C., and Fajkus, J. (2015). Homology-dependent repair is involved in 45S rDNA loss in plant CAF-1 mutants. Plant J *81*, 198-209.

Musielak, T.J., Schenkel, L., Kolb, M., Henschen, A., and Bayer, M. (2015). A simple and versatile cell wall staining protocol to study plant reproduction. Plant Reprod *28*, 161-169.

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y.*, et al.* (2007). Improved Gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. Biosci Biotechnol Biochem *71*, 2095-2100.

Nakajima, K., Sena, G., Nawy, T., and Benfey, P.N. (2001). Intercellular movement of the putative transcription factor SHR in root patterning. Nature *413*, 307-311.

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmler, K., Hause, G., Aurich, A.C., Dorn, R., Jenuwein, T.*, et al.* (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. EMBO J *24*, 1418-1429.

Navashin, M. (1934). Chromosome alterations caused by hybridization and their bearing upon certain general genetic problems. Cytologia *5*, 169-203.

Ng, D.W., Wang, T., Chandrasekharan, M.B., Aramayo, R., Kertbundit, S., and Hall, T.C. (2007). Plant SET domain-containing proteins: structure, function and regulation. Biochim Biophys Acta *1769*, 316-329.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature *463*, 457-463.

Oda, M., Glass, J.L., Thompson, R.F., Mo, Y., Olivier, E.N., Figueroa, M.E., Selzer, R.R., Richmond, T.A., Zhang, X., Dannenberg, L.*, et al.* (2009). High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. Nucleic Acids Res *37*, 3829-3839.

Olins, A.L., Carlson, R.D., and Olins, D.E. (1975). Visualization of chromatin substructure: upsilon bodies. J Cell Biol *64*, 528-537.

Olmedo-Monfil, V., Durán-Figueroa, N., Arteaga-Vázquez, M., Demesa-Arévalo, E., Autran, D., Grimanelli, D., Slotkin, R.K., Martienssen, R.A., and Vielle-Calzada, J.-P. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature *464*, 628-632.

Olson, M.O.J. (2011). The nucleolus. In, M.O.J. Olson, ed. (New York: Springer), pp. V-XVI.

Onate-Sanchez, L., and Vicente-Carbajosa, J. (2008). DNA-free RNA isolation protocols for Arabidopsis thaliana, including seeds and siliques. BMC Res Notes *1*, DOI: 10.1186/1756-0500-1-93.

Onder, O., Sidoli, S., Carroll, M., and Garcia, B.A. (2015). Progress in epigenetic histone modification analysis by mass spectrometry for clinical investigations. Expert Rev Proteomics *12*, 499-517.

Orman-Ligeza, B., Parizot, B., de Rycke, R., Fernandez, A., Himschoot, E., Van Breusegem, F., Bennett, M.J., Perilleux, C., Beeckman, T., and Draye, X. (2016). RBOH-mediated ROS production facilitates lateral root emergence in Arabidopsis. Development *143*, 3328-3339.

Otero, S., Desvoyes, B., Peiro, R., and Gutierrez, C. (2016). Histone H3 dynamics reveal domains with distinct proliferation potential in the Arabidopsis root. Plant Cell *28*, 1361-1371.

Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., and O'Shea, C.C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science *357*.

Pavlistova, V., Dvorackova, M., Jez, M., Mozgova, I., Mokros, P., and Fajkus, J. (2016). Phenotypic reversion in fas mutants of Arabidopsis thaliana by reintroduction of FAS genes: variable recovery of telomeres with major spatial rearrangements and transcriptional reprogramming of 45S rDNA genes. Plant J *88*, 411-424.

Perianez-Rodriguez, J., Manzano, C., and Moreno-Risueno, M.A. (2014). Post-embryonic organogenesis and plant regeneration from tissues: two sides of the same coin? Front Plant Sci *5*, 219.

Petricka, J.J., Winter, C.M., and Benfey, P.N. (2012). Control of Arabidopsis root development. Annu Rev Plant Biol *63*, 563-590.

Pien, S., Fleury, D., Mylne, J.S., Crevillen, P., Inze, D., Avramova, Z., Dean, C., and Grossniklaus, U. (2008). ARABIDOPSIS TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. Plant Cell *20*, 580-588.

Pikaard, C.S. (2003). Nucleolar Dominance. In Encyclopedia of the human genome, D.N. Cooper, ed. (UK: Macmillan Publishers Ltd), pp. 1-5.

Pillot, M., Baroux, C., Vazquez, M.A., Autran, D., Leblanc, O., Vielle-Calzada, J.P., Grossniklaus, U., and Grimanelli, D. (2010). Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in Arabidopsis. Plant Cell *22*, 307-320.

Pontes, O., Lawrence, R.J., Silva, M., Preuss, S., Costa-Nunes, P., Earley, K., Neves, N., Viegas, W., and Pikaard, C.S. (2007). Postembryonic establishment of megabase-scale gene silencing in nucleolar dominance. PLoS One *2*, e1157.

Pontvianne, F., Abou-Ellail, M., Douet, J., Comella, P., Matia, I., Chandrasekhara, C., Debures, A., Blevins, T., Cooke, R., Medina, F.J., *et al.* (2010). Nucleolin is required for DNA methylation state and the expression of rRNA gene variants in Arabidopsis thaliana. PLoS Genet *6*, e1001225.

Pontvianne, F., Blevins, T., Chandrasekhara, C., Feng, W., Stroud, H., Jacobsen, S.E., Michaels, S.D., and Pikaard, C.S. (2012). Histone methyltransferases regulating rRNA gene dose and dosage control in Arabidopsis. Genes Dev *26*, 945-957.

Pontvianne, F., Blevins, T., Chandrasekhara, C., Mozgova, I., Hassel, C., Pontes, O.M., Tucker, S., Mokros, P., Muchova, V., Fajkus, J., *et al.* (2013). Subnuclear partitioning of rRNA genes between the nucleolus and nucleoplasm reflects alternative epiallelic states. Genes Dev *27*, 1545-1550.

Pontvianne, F., Carpentier, M.C., Durut, N., Pavlistova, V., Jaske, K., Schorova, S., Parrinello, H., Rohmer, M., Pikaard, C.S., Fojtova, M., *et al.* (2016). Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the A. thaliana Genome. Cell Rep *16*, 1574-1587.

Preuss, S.B., Costa-Nunes, P., Tucker, S., Pontes, O., Lawrence, R.J., Mosher, R., Kasschau, K.D., Carrington, J.C., Baulcombe, D.C., Viegas, W., *et al.* (2008). Multimegabase silencing in nucleolar dominance involves siRNA-directed DNA methylation and specific methylcytosine-binding proteins. Mol Cell *32*, 673-684.

Prunet, N., and Jack, T. (2014). Flower development in Arabidopsis: there is more to It than learning your ABCs. In Methods Mol Biol, L. Riechmann, and F. Wellmer, eds. (New York: Humana Press), pp. 3-33.

Raatz, B., Eicker, A., Schmitz, G., Fuss, E., Muller, D., Rossmann, S., and Theres, K. (2011). Specific expression of LATERAL SUPPRESSOR is controlled by an evolutionarily conserved 3' enhancer. Plant J *68*, 400-412.

Rabanal, F.A., Mandakova, T., Soto-Jimenez, L.M., Greenhalgh, R., Parrott, D.L., Lutzmayer, S., Steffen, J.G., Nizhynska, V., Mott, R., Lysak, M.A., *et al.* (2017). Epistatic and allelic interactions control expression of ribosomal RNA gene clusters in Arabidopsis thaliana. Genome Biol *18*. DOI: 10.1186/s13059-017-1209-z

Rahman, M.A., Kristiansen, P.E., Veiseth, S.V., Andersen, J.T., Yap, K.L., Zhou, M.M., Sandlie, I., Thorstensen, T., and Aalen, R.B. (2014). The Arabidopsis histone methyltransferase SUVR4 binds ubiquitin via a domain with a four-helix bundle structure. Biochemistry *53*, 2091-2100.

Ramirez-Parra, E., and Gutierrez, C. (2007). E2F regulates FASCIATA1, a chromatin assembly genem whose loss switches on the endocycle and activates gene expression by changing the epigenetic status. Plant Physiol *144*, 105-120.

Raynaud, C., Sozzani, R., Glab, N., Domenichini, S., Perennes, C., Cella, R., Kondorosi, E., and Bergounioux, C. (2006). Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. Plant J *47*, 395-407.

Rivera, C., Gurard-Levin, Z.A., Almouzni, G., and Loyola, A. (2014). Histone lysine methylation and chromatin replication. Biochim Biophys Acta *1839*, 1433-1439.

Robert, H.S., Grunewald, W., Sauer, M., Cannoot, B., Soriano, M., Swarup, R., Weijers, D., Bennett, M., Boutilier, K., and Friml, J. (2015). Plant embryogenesis requires AUX/LAX-mediated auxin influx. Development *142*, 702-711.

Rosas, U., Mei, Y., Xie, Q., Banta, J.A., Zhou, R.W., Seufferheld, G., Gerard, S., Chou, L., Bhambhra, N., Parks, J.D., *et al.* (2014). Variation in Arabidopsis flowering time associated with cis-regulatory variation in CONSTANS. Nat Commun *5*, 3651.

Roudier, F., Ahmed, I., Berard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berthet, E., Al-Shikhley, L., *et al.* (2011). Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. EMBO J *30*, 1928-1938.

Ruckova, E., Friml, J., Prochazkova Schrumpfova, P., and Fajkus, J. (2008). Role of alternative telomere lengthening unmasked in telomerase knockout mutant plants. Plant Mol Biol *66*, 637–646.

Ruda, D., and Warner, J.R. (2004). What better measure than ribosome synthesis? Genes and development *18*, 2431-2436.

Sabatini, S., Heidstra, R., Wildwater, M., and Scheres, B. (2003). SCARECROW is involved in positioning the stem cell niche in the Arabidopsis root meristem. Genes and Development *17*, 354-358.

Sahr, T., Adam, T., Fizames, C., Maurel, C., and Santoni, V. (2010). O-carboxyl- and N-methyltransferases active on plant aquaporins. Plant Cell Physiol *51*, 2092-2104.

Saleh, A., Al-Abdallat, A., Ndamukong, I., Alvarez-Venegas, R., and Avramova, Z. (2007). The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus. Nucleic Acids Res *35*, 6290-6296.

Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Le, O., Hou, G., Sadder, M., Al-Abdallat, A., Xia, Y., Lu, G., Ladunga, I.*, et al.* (2008). The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. Plant Cell *20*, 568-579.

Sanchez, M.d.L., Costas, C., Sequeira-Mendes, J., and Gutierrez, C. (2012). Regulating DNA replication in plants. Cold Spring Harb Perspect Biol *4*, a010140.

Sanchez, M.d.L., and Gutierrez, C. (2009a). Arabidopsis ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. Proc Natl Acad Sci U S A *106*, 2065-2070.

Sanchez, M.d.L., and Gutierrez, C. (2009b). Novel insights into the plant histone code: Lessons from ORC1. Epigenetics *4*, 205-208.

Sarkar, A.K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., Scheres, B., Heidstra, R., and Laux, T. (2007). Conserved factors regulate signalling in Arabidopsis thaliana shoot and root stem cell organizers. Nature *446*, 811-814.

Schneitz, K., Hülskamp, M., and Pruitt, R. (1995). Wild-type ovule development in Arabidopsis thaliana: a light microscope study of cleared whole-mount tissue. Plant J *7*, 731-739.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev *18*, 1251-1262.

Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A.*, et al.* (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev *22*, 2048-2061.

Seeliger, K., Dukowic-Schulze, S., Wurz-Wildersinn, R., Pacher, M., and Puchta, H. (2012). BRCA2 is a mediator of RAD51- and DMC1-facilitated homologous recombination in Arabidopsis thaliana. New Phytol *193*, 364-375.

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The functional topography of the Arabidopsis genome is organized in a reduced number of linear motifs of chromatin states. Plant Cell *26*, 2351-2366.

Sequeira-Mendes, J., and Gutierrez, C. (2015). Links between genome replication and chromatin landscapes. Plant J *83*, 38-51.

Sequeira-Mendes, J., and Gutierrez, C. (2016). Genome architecture: from linear organisation of chromatin to the 3D assembly in the nucleus. Chromosoma *125*, 455-469.

Serrano-Mislata, A., Fernandez-Nohales, P., Domenech, M.J., Hanzawa, Y., Bradley, D., and Madueno, F. (2016). Separate elements of the TERMINAL FLOWER 1 cis-regulatory region integrate pathways to control flowering time and shoot meristem identity. Development *143*, 3315-3327.

Shafiq, S., Berr, A., and Shen, W.H. (2014). Combinatorial functions of diverse histone methylations in Arabidopsis thaliana flowering time regulation. New Phytol *201*, 312-322.

Shaw, P., and Brown, J. (2012). Nucleoli: composition, function, and dynamics. Plant Physiol *158*, 44-51.

Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. Nat Protoc *2*, 1445-1457.

Shubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. EMBO J *25*, 4638-4649.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461-472.

Smyth, D.R., Bowman, J.L., and Meyerowitz, E.M. (1990). Early flower development in Arabídopsis. Plant Cell *2*, 755-766.

Song, J., Keppler, B.D., Wise, R.R., and Bent, A.F. (2015). PARP2 Is the Predominant Poly(ADP-Ribose) Polymerase in Arabidopsis DNA Damage and Immune Responses. PLoS Genet *11*, e1005200.

Sparks, E., Wachsman, G., and Benfey, P.N. (2013). Spatiotemporal signalling in plant development. Nat Rev Genet *14*, 631-644.

Spencer, M.W., Casson, S.A., and Lindsey, K. (2007). Transcriptional profiling of the Arabidopsis embryo. Plant Physiol *143*, 924-940.

Spillane, C., MacDougall, C., Stock, C., Köhler, C., Vielle-Calzada, J.P., Nunes, S.M., Grossniklaus, U., and Goodrich, J. (2000). Interaction of the Arabidopsis Polycomb group proteins FIE and MEA mediates their common phenotypes. Curr Biol *10*, 1535-1538.

Stepinski, D. (2010). Organization of the nucleoli of soybean root meristematic cells at different states of their activity. Micron *41*, 283-288.

Stepinski, D. (2014). Functional ultrastructure of the plant nucleolus. Protoplasma *251*, 1285-1306.

Sundaresan, V., and Alandete-Saez, M. (2010). Pattern formation in miniature: the female gametophyte of flowering plants. Development *137*, 179-189.

Tamada, Y., Yun, J.Y., Woo, S.C., and Amasino, R.M. (2009). ARABIDOPSIS TRITHORAX-RELATED7 is required for methylation of lysine 4 of histone H3 and for transcriptional activation of FLOWERING LOCUS C. Plant Cell *21*, 3257-3269.

Tang, L., Nogales, E., and Ciferri, C. (2010). Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. Prog Biophys Mol Biol *102*, 122-128.

Thorstensen, T., Fischer, A., Sandvik, S.V., Johnsen, S.S., Grini, P.E., Reuter, G., and Aalen, R.B. (2006). The Arabidopsis SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. Nucleic Acids Res *34*, 5461-5470.

Thorstensen, T., Grini, P.E., and Aalen, R.B. (2011). SET domain proteins in plant development. Biochim Biophys Acta *1809*, 407-420.

Thorstensen, T., Grini, P.E., Mercy, I.S., Alm, V., Erdal, S., Aasland, R., and Aalen, R.B. (2008). The Arabidopsis SET-domain protein ASHR3 is involved in stamen development and interacts with the bHLH transcription factor ABORTED MICROSPORES (AMS). Plant Mol Biol *66*, 47-59.

Tominaga-Wada, R., Ishida, T., and Wada, T. (2011). New insights into the mechanism of development of Arabidopsis root hairs and trichomes. Int Rev Cell Mol Biol *286*, 67-106.

Trejo-Arellano, M.S., Mahrez, W., Nakamura, M., Moreno-Romero, J., Nanni, P., Kohler, C., and Hennig, L. (2017). H3K23me1 is an evolutionarily conserved histone modification associated with CG DNA methylation in Arabidopsis. Plant J *90*, 293-303.

Tsang, L.W., Hu, N., and Underhill, D.A. (2010). Comparative analyses of SUV420H1 isoforms and SUV420H2 reveal differences in their cellular localization and effects on myogenic differentiation. PLoS One *5*, e14447.

Tucker, S., Vitins, A., and Pikaard, C.S. (2010). Nucleolar dominance and ribosomal RNA gene silencing. Curr Opin Cell Biol *22*, 351-356.

Ubeda-Tomas, S., Federici, F., Casimiro, I., Beemster, G.T., Bhalerao, R., Swarup, R., Doerner, P., Haseloff, J., and Bennett, M.J. (2009). Gibberellin signaling in the endodermis controls Arabidopsis root meristem size. Curr Biol *19*, 1194-1199.

Valencia-Morales, M.d.P., Camas-Reyes, J.A., Cabrera-Ponce, J.L., and Alvarez-Venegas, R. (2012). The Arabidopsis thaliana SET-domain-containing protein ASHH1/SDG26 interacts with itself and with distinct histone lysine methyltransferases. J Plant Res *125*, 679-692.

Van Norman, J.M., and Benfey, P.N. (2009). Arabidopsis thaliana as a model organism in systems biology. Wiley Interdiscip Rev Syst Biol Med *1*, 372-379.

van Sluis, M., and McStay, B. (2015). A localized nucleolar DNA damage response facilitates recruitment of the homology-directed repair machinery independent of cell cycle stage. Genes Dev *29*, 1151-1163.

van Sluis, M., and McStay, B. (2017). Nucleolar reorganization in response to rDNA damage. Curr Opin Cell Biol *46*, 81-86.

Vanhaeren, H., Gonzalez, N., and Inze, D. (2015). A journey through a leaf: phenomics analysis of leaf growth in Arabidopsis thaliana. Arabidopsis Book *13*, e0181.

Veiseth, S.V., Rahman, M.A., Yap, K.L., Fischer, A., Egge-Jacobsen, W., Reuter, G., Zhou, M.M., Aalen, R.B., and Thorstensen, T. (2011). The SUVR4 histone lysine methyltransferase binds ubiquitin and converts H3K9me1 to H3K9me3 on transposon chromatin in Arabidopsis. PLoS Genet *7*, e1001325.

Verbelen, J.P., De Cnodder, T., Le, J., Vissenberg, K., and Baluska, F. (2006). The root apex of Arabidopsis thaliana consists of four distinct zones of growth activities: meristematic zone, transition zone, fast elongation zone and growth terminating zone. Plant Signal Behav *1*, 296-304.

Vergara, Z. (2017). Cellular and Developmental Control of Arabidopsis DNA Replication: Function of ORC1 and Transposons. Ph.D dissertation. Universidad Autonoma de Madrid. Madrid-Spain

Vergara, Z., and Gutierrez, C. (2017). Emerging roles of chromatin in the maintenance of genome organization and function in plants. Genome Biol *18*, DOI: 10.1186/s13059-017-1236-9.

Vermeer, J.E., and Geldner, N. (2015). Lateral root initiation in Arabidopsis thaliana: a force awakens. F1000Prime Rep *7*, 7-32.

Villar, C.B., and Kohler, C. (2010). Plant chromatin immunoprecipitation. Methods Mol Biol *655*, 401-411.

Wang, D., Tyson, M.D., Jackson, S.S., and Yadegari, R. (2006). Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis. Proc Natl Acad Sci U S A *103*, 13244-13249.

Wang, R., Kamgoue, A., Normand, C., Leger-Silvestre, I., Mangeat, T., and Gadal, O. (2016). High resolution microscopy reveals the nuclear shape of budding yeast during cell cycle and in various biological states. J Cell Sci *129*, 4480-4495.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q.*, et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet *40*, 897-903.

Warmerdam, D.O., van den Berg, J., and Medema, R.H. (2016). Breaks in the 45S rDNA lead to recombination-mediated loss of repeats. Cell Rep *14*, 2519-2527.

References

Weirich, S., Kudithipudi, S., and Jeltsch, A. (2016). Specificity of the SUV4-20H1 and SUV4-20H2 protein lysine methyltransferases and methylation of novel substrates. J Mol Biol *428*, 2344-2358.

West, C.E., Waterworth, W.M., Sunderland, P.A., and Bray, C.M. (2004). Arabidopsis DNA double-strand break repair pathways. Biochem Soc Trans *32*, 964-966.

Wilson, M., Goh, T., Voss, U., Bishopp, A., Peret, B., and Bennett, M. (2013). SnapShot: Root development. Cell *155*, 1190-1191.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS One *2*, e718.

Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J.P., Steinmetz, A., and Shen, W.H. (2008). Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in Arabidopsis thaliana. Mol Cell Biol *28*, 1348-1360.

Yadegari, R., and Drews, G.N. (2004). Female gametophyte development. Plant Cell *16 Suppl*, S133-141.

Yao, X., Feng, H., Yu, Y., Dong, A., and Shen, W.H. (2013). SDG2-mediated H3K4 methylation is required for proper Arabidopsis root growth and development. PLoS One *8*, e56537.

Ying, Z., Janney, N., and Houtz , R. (1996). Organization and characterization of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit epsilon N-methyltransferase gene in tobacco. Plant Mol Biol *32*, 663-671.

Yu, C.W., Tai, R., Wang, S.C., Yang, P., Luo, M., Yang, S., Cheng, K., Wang, W.C., Cheng, Y.S., and Wu, K. (2017). HISTONE DEACETYLASE6 acts in concert with histone methyltransferases SUVH4, SUVH5, and SUVH6 to regulate transposon silencing. Plant Cell *29*, 1970-1983.

Yun, J.Y., Tamada, Y., Kang, Y.E., and Amasino, R.M. (2012). Arabidopsis trithorax-related3/SET domain GROUP2 is required for the winter-annual habit of Arabidopsis thaliana. Plant Cell Physiol *53*, 834-846.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell *153*, 193-205.

Zhang, K., Sridhar, V.V., Zhu, J., Kapoor, A., and Zhu, J.K. (2007). Distinctive core histone post-translational modification patterns in Arabidopsis thaliana. PLoS One *2*, e1210.

Zhang, L., and Ma, H. (2012). Complex evolutionary history and diverse domain organization of SET proteins suggest divergent regulatory interactions. New Phytol *195*, 248-263.

Zhang, W., Swarup, R., Bennett, M., Schaller, G.E., and Kieber, J.J. (2013). Cytokinin induces cell division in the quiescent center of the Arabidopsis root apical meristem. Curr Biol *23*, 1979-1989.

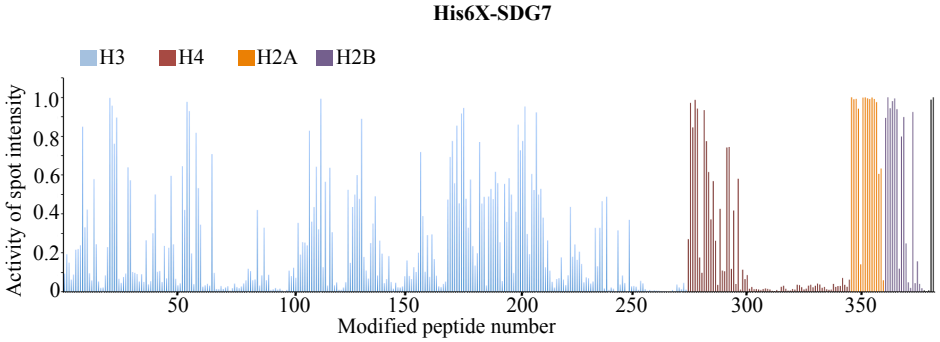
Zhao, W., Shafiq, S., Berr, A., and Shen, W.H. (2015). Genome-wide gene expression profiling to investigate molecular phenotypes of Arabidopsis mutants deprived in distinct histone methyltransferases and demethylases. Genom Data *4*, 143-145.

Zhao, Z., Tuakli-Wosornu, Y., Lagace, T.A., Kinch, L., Grishin, N.V., Horton, J.D., Cohen, J.C., and Hobbs, H.H. (2006). Molecular characterization of loss-of-function mutations in PCSK9 and identification of a compound heterozygote. Am J Hum Genet *79*, 514-523.

Appendix

Appendix 1. Histone peptide binding results

The peptide binding arrays incubated with His6X-SDG7 and His6X-SDG24.1 were analyzed by Western blot against the His6X tag and quantified with the aid of the array analysis software (Active Motif). Here is presented the average of activity of spot intensity obtained for all 384 peptide combinations and their respective locations in the peptide.



Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
1	A1	H3 1-19	unmod				0.094
2	A2	H3 1-19	R2me2s				0.192
3	A3	H3 1-19	R2me2a				0.149
4	A4	H3 1-19	R2Citr				0.062
5	A5	H3 1-19	T3P				0.088
6	A6	H3 1-19	K4me1				0.216
7	A7	H3 1-19	K4me2				0.219
8	A8	H3 1-19	K4me3				0.239
9	A9	H3 1-19	K4ac				0.849
10	A10	H3 1-19	R8me2s				0.331
11	A11	H3 1-19	R8me2a				0.422
12	A12	H3 1-19	R8Citr				0.095
13	A13	H3 1-19	K9me1				0.058
14	A14	H3 1-19	K9me2				0.579
15	A15	H3 1-19	K9me3				0.244
16	A16	H3 1-19	K9ac				0.052
17	A17	H3 1-19	S10P				0.020
18	A18	H3 1-19	T11P				0.022
19	A19	H3 1-19	K14ac				0.084
20	A20	H3 1-19	R2me2s	T3P			0.229
21	A21	H3 1-19	R2me2s	K4me1			0.997
22	A22	H3 1-19	R2me2s	K4me2			0.957
23	A23	H3 1-19	R2me2s	K4me3			0.761
24	A24	H3 1-19	R2me2s	K4ac			0.896
25	B1	H3 1-19	R2me2a	T3P			0.066
26	B2	H3 1-19	R2me2a	K4me1			0.045
27	B3	H3 1-19	R2me2a	K4me2			0.075
28	B4	H3 1-19	R2me2a	K4me3			0.093
29	B5	H3 1-19	R2me2a	K4ac			0.640
30	B6	H3 1-19	R2Citr	T3P			0.573
31	B7	H3 1-19	R2Citr	K4me1			0.101
32	B8	H3 1-19	R2Citr	K4me2			0.097
33	B9	H3 1-19	R2Citr	K4me3			0.091
34	B10	H3 1-19	R2Citr	K4ac			0.052
35	B11	H3 1-19	T3P	K4me1			0.089
36	B12	H3 1-19	T3P	K4me2			0.050
37	B13	H3 1-19	T3P	K4me3			0.264
38	B14	H3 1-19	T3P	K4ac			0.038
39	B15	H3 1-19	R2me2s	T3P	K4me1		0.054
40	B16	H3 1-19	R2me2s	T3P	K4me2		0.300
41	B17	H3 1-19	R2me2s	T3P	K4me3		0.500
42	B18	H3 1-19	R2me2s	T3P	K4ac		0.102
43	B19	H3 1-19	R2me2a	T3P	K4me1		0.106
44	B20	H3 1-19	R2me2a	T3P	K4me2		0.050
45	B21	H3 1-19	R2me2a	T3P	K4me3		0.236
46	B22	H3 1-19	R2me2a	T3P	K4ac		0.068
47	B23	H3 1-19	R8me2s	K9me1			0.226
48	B24	H3 1-19	R8me2s	K9me2			0.596

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
49	C1	H3 1-19	R8me2s	K9me3			0.243
50	C2	H3 1-19	R8me2s	K9ac			0.066
51	C3	H3 1-19	R8me2s	S10P			0.033
52	C4	H3 1-19	R8me2s	T11P			0.043
53	C5	H3 1-19	R8me2a	K9me1			0.645
54	C6	H3 1-19	R8me2a	K9me2			0.420
55	C7	H3 1-19	R8me2a	K9me3			0.977
56	C8	H3 1-19	R8me2a	K9ac			0.929
57	C9	H3 1-19	R8me2a	S10P			0.197
58	C10	H3 1-19	R8me2a	T11P			0.039
59	C11	H3 1-19	R8Citr	K9me1			0.817
60	C12	H3 1-19	R8Citr	K9me2			0.533
61	C13	H3 1-19	R8Citr	K9me3			0.345
62	C14	H3 1-19	R8Citr	K9ac			0.026
63	C15	H3 1-19	R8Citr	S10P			0.032
64	C16	H3 1-19	R8Citr	T11P			0.040
65	C17	H3 1-19	K9me1	S10P			0.031
66	C18	H3 1-19	K9me1	T11P			0.708
67	C19	H3 1-19	K9me1	K14ac			0.096
68	C20	H3 1-19	K9me2	S10P			0.014
69	C21	H3 1-19	K9me2	T11P			0.013
70	C22	H3 1-19	K9me2	K14ac			0.029
71	C23	H3 1-19	K9me3	S10P			0.014
72	C24	H3 1-19	K9me3	T11P			0.021
73	D1	H3 1-19	K9me3	K14ac			0.028
74	D2	H3 1-19	K9ac	S10P			0.005
75	D3	H3 1-19	K9ac	T11P			0.018
76	D4	H3 1-19	K9ac	K14ac			0.042
77	D5	H3 1-19	S10P	T11P			0.023
78	D6	H3 1-19	S10P	K14ac			0.018
79	D7	H3 1-19	T11P	K14ac			0.027
80	D8	H3 1-19	R8me2s	K9me1	S10P		0.042
81	D9	H3 1-19	R8me2s	K9me2	S10P		0.058
82	D10	H3 1-19	R8me2s	K9me3	S10P		0.118
83	D11	H3 1-19	R8me2s	K9ac	S10P		0.106
84	D12	H3 1-19	R8me2s	K9me1	T11P		0.059
85	D13	H3 1-19	R8me2s	K9me2	T11P		0.063
86	D14	H3 1-19	R8me2s	K9me3	T11P		0.420
87	D15	H3 1-19	R8me2s	K9ac	T11P		0.031
88	D16	H3 1-19	R8me2a	K9me1	S10P		0.084
89	D17	H3 1-19	R8me2a	K9me2	S10P		0.329
90	D18	H3 1-19	R8me2a	K9me3	S10P		0.059
91	D19	H3 1-19	R8me2a	K9ac	S10P		0.087
92	D20	H3 1-19	R8me2a	K9me1	T11P		0.008
93	D21	H3 1-19	R8me2a	K9me2	T11P		0.012
94	D22	H3 1-19	R8me2a	K9me3	T11P		0.019
95	D23	H3 1-19	R8me2a	K9ac	T11P		0.011
96	D24	H3 1-19	R8me2a	K9me1	S10P	T11P	0.015
97	E1	H3 1-19	R8me2a	K9me2	S10P	T11P	0.006
98	E2	H3 1-19	R8me2a	K9me3	S10P	T11P	0.000
99	E3	H3 1-19	R8me2a	K9ac	S10P	T11P	0.009
100	E4	H3 1-19	R2me2s	K4me1	R8me2s		0.108
101	E5	H3 1-19	R2me2s	K4me2	R8me2s		0.080
102	E6	H3 1-19	R2me2s	K4me3	R8me2s		0.123
103	E7	H3 1-19	R2me2s	K4ac	R8me2s		0.071
104	E8	H3 1-19	R2me2a	K4me1	R8me2a		0.354
105	E9	H3 1-19	R2me2a	K4me2	R8me2a		0.191
106	E10	H3 1-19	R2me2a	K4me3	R8me2a		0.255
107	E11	H3 1-19	R2me2a	K4ac	R8me2a		0.252
108	E12	H3 1-19	R2me2s	K4me1	K9me1		0.238
109	E13	H3 1-19	R2me2s	K4me2	K9me1		0.828
110	E14	H3 1-19	R2me2s	K4me3	K9me1		0.360
111	E15	H3 1-19	R2me2s	K4ac	K9me1		0.435
112	E16	H3 1-19	R2me2a	K4me1	K9me2		0.643
113	E17	H3 1-19	R2me2a	K4me2	K9me2		0.321
114	E18	H3 1-19	R2me2a	K4me3	K9me2		0.993
115	E19	H3 1-19	R2me2a	K4ac	K9me2		0.126
116	E20	H3 1-19	R2me2s	K4me1	K9me3		0.565
117	E21	H3 1-19	R2me2s	K4me2	K9me3		0.032
118	E22	H3 1-19	R2me2s	K4me3	K9me3		0.637
119	E23	H3 1-19	R2me2s	K4ac	K9me3		0.306
120	E24	H3 1-19	R2me2a	K4me1	K9ac		0.032
121	F1	H3 1-19	R2me2a	K4me2	K9ac		0.048
122	F2	H3 1-19	R2me2a	K4me3	K9ac		0.010
123	F3	H3 1-19	R2me2a	K4ac	K9ac		0.010
124	F4	H3 1-19	K4me1	R8me2s	K9me1		0.021

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
125	F5	H3 1-19	K4me2	R8me2s	K9me1		0.049
126	F6	H3 1-19	K4me3	R8me2s	K9me1		0.524
127	F7	H3 1-19	K4ac	R8me2s	K9me1		0.148
128	F8	H3 1-19	K4me1	R8me2a	K9me1		0.436
129	F9	H3 1-19	K4me2	R8me2a	K9me1		0.500
130	F10	H3 1-19	K4me3	R8me2a	K9me1		0.600
131	F11	H3 1-19	K4ac	R8me2a	K9me1		0.477
132	F12	H3 1-19	K4me1	R8me2s	K9me2		0.889
133	F13	H3 1-19	K4me2	R8me2s	K9me2		0.178
134	F14	H3 1-19	K4me3	R8me2s	K9me2		0.109
135	F15	H3 1-19	K4ac	R8me2s	K9me2		0.067
136	F16	H3 1-19	K4me1	R8me2a	K9me2		0.250
137	F17	H3 1-19	K4me2	R8me2a	K9me2		0.351
138	F18	H3 1-19	K4me3	R8me2a	K9me2		0.491
139	F19	H3 1-19	K4ac	R8me2a	K9me2		0.083
140	F20	H3 1-19	K4me1	R8me2s	K9me3		0.263
141	F21	H3 1-19	K4me2	R8me2s	K9me3		0.195
142	F22	H3 1-19	K4me3	R8me2s	K9me3		0.044
143	F23	H3 1-19	K4ac	R8me2s	K9me3		0.065
144	F24	H3 1-19	K4me1	R8me2a	K9me3		0.183
145	G1	H3 1-19	K4me2	R8me2a	K9me3		0.050
146	G2	H3 1-19	K4me3	R8me2a	K9me3		0.015
147	G3	H3 1-19	K4ac	R8me2a	K9me3		0.046
148	G4	H3 1-19	K4me1	R8me2s	K9ac		0.023
149	G5	H3 1-19	K4me2	R8me2s	K9ac		0.022
150	G6	H3 1-19	K4me3	R8me2s	K9ac		0.027
151	G7	H3 1-19	K4ac	R8me2s	K9ac		0.080
152	G8	H3 1-19	K4me1	R8me2a	K9ac		0.160
153	G9	H3 1-19	K4me2	R8me2a	K9ac		0.081
154	G10	H3 1-19	K4me3	R8me2a	K9ac		0.065
155	G11	H3 1-19	K4ac	R8me2a	K9ac		0.126
156	G12	H3 1-19	R2me2s	K4me1	R8me2s	K9me1	0.102
157	G13	H3 1-19	R2me2s	K4me2	R8me2s	K9me1	0.200
158	G14	H3 1-19	R2me2s	K4me3	R8me2s	K9me1	0.718
159	G15	H3 1-19	R2me2s	K4ac	R8me2s	K9me1	0.389
160	G16	H3 1-19	R2me2a	K4me1	R8me2s	K9me1	0.102
161	G17	H3 1-19	R2me2a	K4me2	R8me2s	K9me1	0.291
162	G18	H3 1-19	R2me2a	K4me3	R8me2s	K9me1	0.076
163	G19	H3 1-19	R2me2a	K4ac	R8me2s	K9me1	0.295
164	G20	H3 1-19	R2me2s	K4me1	R8me2s	K9me2	0.168
165	G21	H3 1-19	R2me2s	K4me2	R8me2s	K9me2	0.082
166	G22	H3 1-19	R2me2s	K4me3	R8me2s	K9me2	0.027
167	G23	H3 1-19	R2me2s	K4ac	R8me2s	K9me2	0.024
168	G24	H3 1-19	R2me2a	K4me1	R8me2s	K9me2	0.040
169	H1	H3 1-19	R2me2a	K4me2	R8me2s	K9me2	0.052
170	H2	H3 1-19	R2me2a	K4me3	R8me2s	K9me2	0.474
171	H3	H3 1-19	R2me2a	K4ac	R8me2s	K9me2	0.693
172	H4	H3 1-19	R2me2s	K4me1	R8me2s	K9me3	0.775
173	H5	H3 1-19	R2me2s	K4me2	R8me2s	K9me3	0.558
174	H6	H3 1-19	R2me2s	K4me3	R8me2s	K9me3	0.854
175	H7	H3 1-19	R2me2s	K4ac	R8me2s	K9me3	0.454
176	H8	H3 1-19	R2me2a	K4me1	R8me2s	K9me3	0.916
177	H9	H3 1-19	R2me2a	K4me2	R8me2s	K9me3	0.946
178	H10	H3 1-19	R2me2a	K4me3	R8me2s	K9me3	0.478
179	H11	H3 1-19	R2me2a	K4ac	R8me2s	K9me3	0.329
180	H12	H3 1-19	R2me2s	K4me1	R8me2s	K9ac	0.577
181	H13	H3 1-19	R2me2s	K4me2	R8me2s	K9ac	0.234
182	H14	H3 1-19	R2me2s	K4me3	R8me2s	K9ac	0.135
183	H15	H3 1-19	R2me2s	K4ac	R8me2s	K9ac	0.198
184	H16	H3 1-19	R2me2a	K4me1	R8me2s	K9ac	0.770
185	H17	H3 1-19	R2me2a	K4me2	R8me2s	K9ac	0.454
186	H18	H3 1-19	R2me2a	K4me3	R8me2s	K9ac	0.488
187	H19	H3 1-19	R2me2a	K4ac	R8me2s	K9ac	0.032
188	H20	H3 1-19	R2me2s	K4me1	R8me2a	K9me1	0.490
189	H21	H3 1-19	R2me2s	K4me2	R8me2a	K9me1	0.528
190	H22	H3 1-19	R2me2s	K4me3	R8me2a	K9me1	0.476
191	H23	H3 1-19	R2me2s	K4ac	R8me2a	K9me1	0.617
192	H24	H3 1-19	R2me2a	K4me1	R8me2a	K9me1	0.558
193	I1	H3 1-19	R2me2a	K4me2	R8me2a	K9me1	0.254
194	I2	H3 1-19	R2me2a	K4me3	R8me2a	K9me1	0.054
195	I3	H3 1-19	R2me2a	K4ac	R8me2a	K9me1	0.554
196	I4	H3 1-19	R2me2s	K4me1	R8me2a	K9me2	0.359
197	I5	H3 1-19	R2me2s	K4me2	R8me2a	K9me2	0.584
198	I6	H3 1-19	R2me2s	K4me3	R8me2a	K9me2	0.500
199	I7	H3 1-19	R2me2s	K4ac	R8me2a	K9me2	0.096
200	I8	H3 1-19	R2me2a	K4me1	R8me2a	K9me2	0.411

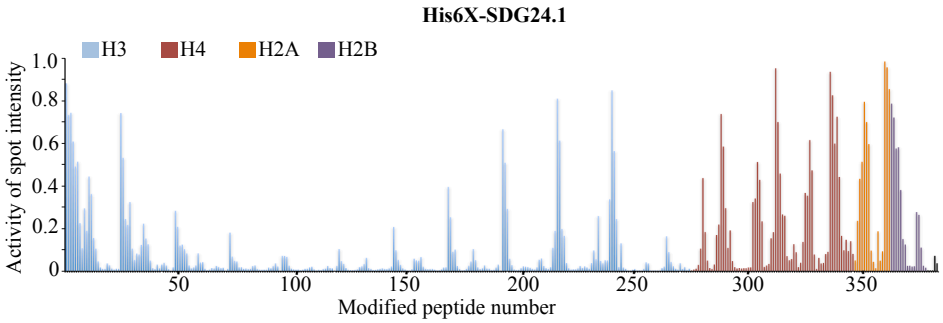
Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
201	I9	H3 1-19	R2me2a	K4me2	R8me2a	K9me2	0.859
202	I10	H3 1-19	R2me2a	K4me3	R8me2a	K9me2	0.728
203	I11	H3 1-19	R2me2a	K4ac	R8me2a	K9me2	0.776
204	I12	H3 1-19	R2me2s	K4me1	R8me2a	K9me3	0.953
205	I13	H3 1-19	R2me2s	K4me2	R8me2a	K9me3	0.297
206	I14	H3 1-19	R2me2s	K4me3	R8me2a	K9me3	0.193
207	I15	H3 1-19	R2me2s	K4ac	R8me2a	K9me3	0.606
208	I16	H3 1-19	R2me2a	K4me1	R8me2a	K9me3	0.523
209	I17	H3 1-19	R2me2a	K4me2	R8me2a	K9me3	0.923
210	I18	H3 1-19	R2me2a	K4me3	R8me2a	K9me3	0.500
211	I19	H3 1-19	R2me2a	K4ac	R8me2a	K9me3	0.529
212	I20	H3 1-19	R2me2s	K4me1	R8me2a	K9ac	0.381
213	I21	H3 1-19	R2me2s	K4me2	R8me2a	K9ac	0.125
214	I22	H3 1-19	R2me2s	K4me3	R8me2a	K9ac	0.264
215	I23	H3 1-19	R2me2s	K4ac	R8me2a	K9ac	0.108
216	I24	H3 1-19	R2me2a	K4me1	R8me2a	K9ac	0.051
217	J1	H3 1-19	R2me2a	K4me2	R8me2a	K9ac	0.021
218	J2	H3 1-19	R2me2a	K4me3	R8me2a	K9ac	0.065
219	J3	H3 1-19	R2me2a	K4ac	R8me2a	K9ac	0.069
220	J4	H3 7-26	unmod				0.176
221	J5	H3 7-26	K14ac				0.062
222	J6	H3 7-26	K14ac	S10P			0.034
223	J7	H3 7-26	K14ac	T11P			0.084
224	J8	H3 7-26	R17me2s				0.436
225	J9	H3 7-26	R17me2a				0.180
226	J10	H3 7-26	R17Citr				0.243
227	J11	H3 7-26	K18ac				0.163
228	J12	H3 7-26	K14ac	R17me2s			0.206
229	J13	H3 7-26	K14ac	R17me2a			0.142
230	J14	H3 7-26	K14ac	K18ac			0.048
231	J15	H3 7-26	R17me2s	K18ac			0.074
232	J16	H3 7-26	R17me2a	K18ac			0.069
233	J17	H3 7-26	R17Citr	K18ac			0.041
234	J18	H3 7-26	K14ac	R17me2s	K18ac		0.052
235	J19	H3 7-26	K14ac	R17me2a	K18ac		0.328
236	J20	H3 16-35	unmod				0.131
237	J21	H3 16-35	R26me2s				0.328
238	J22	H3 16-35	R26me2a				0.466
239	J23	H3 16-35	R26Citr				0.017
240	J24	H3 16-35	K27me1				0.488
241	K1	H3 16-35	K27me2				0.009
242	K2	H3 16-35	K27me3				0.022
243	K3	H3 16-35	K27ac				0.015
244	K4	H3 16-35	S28P				0.002
245	K5	H3 16-35	R26me2s	K27me1			0.315
246	K6	H3 16-35	R26me2s	K27me2			0.037
247	K7	H3 16-35	R26me2s	K27me3			0.083
248	K8	H3 16-35	R26me2s	K27ac			0.044
249	K9	H3 16-35	R26me2s	S28P			0.002
250	K10	H3 16-35	R26me2a	K27me1			0.370
251	K11	H3 16-35	R26me2a	K27me2			0.027
252	K12	H3 16-35	R26me2a	K27me3			0.029
253	K13	H3 16-35	R26me2a	K27ac			0.026
254	K14	H3 16-35	R26me2a	S28P			0.010
255	K15	H3 16-35	R26Citr	K27me1			0.056
256	K16	H3 16-35	R26Citr	K27me2			0.043
257	K17	H3 16-35	R26Citr	K27me3			0.014
258	K18	H3 16-35	R26Citr	S28P			0.005
259	K19	H3 16-35	K27me1	S28P			0.008
260	K20	H3 16-35	K27me2	S28P			0.005
261	K21	H3 16-35	K27me3	S28P			0.007
262	K22	H3 16-35	K27ac	S28P			0.005
263	K23	H3 16-35	R26me2s	K27me1	S28P		0.005
264	K24	H3 16-35	R26me2s	K27me2	S28P		0.003
265	L1	H3 16-35	R26me2s	K27me3	S28P		0.002
266	L2	H3 16-35	R26me2s	K27ac	S28P		0.001
267	L3	H3 16-35	R26me2a	K27me1	S28P		0.000
268	L4	H3 16-35	R26me2a	K27me2	S28P		0.002
269	L5	H3 16-35	R26me2a	K27me3	S28P		0.008
270	L6	H3 16-35	R26me2a	K27ac	S28P		0.003
271	L7	H3 26-45	unmod				0.037
272	L8	H3 26-45	K36me1				0.022
273	L9	H3 26-45	K36me2				0.013
274	L10	H3 26-45	K36me3				0.046
275	L11	H3 26-45	K36ac				0.011
276	L12	H4 1-19	unmod				0.270

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
277	L13	H4 1-19	S1P				0.972
278	L14	H4 1-19	R3me2s				0.845
279	L15	H4 1-19	R3me2a				0.987
280	L16	H4 1-19	K5ac				0.942
281	L17	H4 1-19	K8ac				0.176
282	L18	H4 1-19	K12ac				0.097
283	L19	H4 1-19	K16ac				0.934
284	L20	H4 1-19	S1P	R3me2s			0.774
285	L21	H4 1-19	S1P	R3me2a			0.616
286	L22	H4 1-19	S1P	K5ac			0.372
287	L23	H4 1-19	R3me2s	K5ac			0.569
288	L24	H4 1-19	R3me2s	K8ac			0.262
289	M1	H4 1-19	R3me2a	K5ac			0.035
290	M2	H4 1-19	R3me2a	K8ac			0.426
291	M3	H4 1-19	K5ac	K8ac			0.109
292	M4	H4 1-19	K8ac	K12ac			0.105
293	M5	H4 1-19	K8ac	K16ac			0.742
294	M6	H4 1-19	K12ac	K16ac			0.745
295	M7	H4 1-19	S1P	R3me2s	K5ac		0.117
296	M8	H4 1-19	S1P	R3me2a	K5ac		0.418
297	M9	H4 1-19	R3me2s	K5ac	K8ac		0.040
298	M10	H4 1-19	R3me2a	K5ac	K8ac		0.581
299	M11	H4 1-19	K5ac	K8ac	K12ac		0.011
300	M12	H4 1-19	K8ac	K12ac	K16ac		0.113
301	M13	H4 1-19	S1P	R3me2s	K5ac	K8ac	0.065
302	M14	H4 1-19	S1P	R3me2a	K5ac	K8ac	0.085
303	M15	H4 1-19	R3me2s	K5ac	K8ac	K12ac	0.015
304	M16	H4 1-19	R3me2a	K5ac	K8ac	K12ac	0.024
305	M17	H4 1-19	K5ac	K8ac	K12ac	K16ac	0.013
306	M18	H4 11-30	unmod				0.013
307	M19	H4 11-30	K12ac				0.011
308	M20	H4 11-30	K16ac				0.009
309	M21	H4 11-30	R17me2s				0.014
310	M22	H4 11-30	R17me2a				0.016
311	M23	H4 11-30	R19me2s				0.015
312	M24	H4 11-30	R19me2a				0.012
313	N1	H4 11-30	K20me1				0.005
314	N2	H4 11-30	K20me2				0.003
315	N3	H4 11-30	K20me3				0.008
316	N4	H4 11-30	K20ac				0.004
317	N5	H4 11-30	R24me2a				0.025
318	N6	H4 11-30	R24me2s				0.031
319	N7	H4 11-30	K12ac	K16ac			0.015
320	N8	H4 11-30	K16ac	R17me2s			0.005
321	N9	H4 11-30	K16ac	R17me2a			0.006
322	N10	H4 11-30	K16ac	R19me2s			0.014
323	N11	H4 11-30	K16ac	R19me2a			0.011
324	N12	H4 11-30	K16ac	K20me1			0.036
325	N13	H4 11-30	K16ac	K20me2			0.034
326	N14	H4 11-30	K16ac	K20me3			0.023
327	N15	H4 11-30	K16ac	K20ac			0.011
328	N16	H4 11-30	K12ac	K16ac	K20me1		0.016
329	N17	H4 11-30	K12ac	K16ac	K20me2		0.031
330	N18	H4 11-30	K12ac	K16ac	K20me3		0.036
331	N19	H4 11-30	K12ac	K16ac	K20ac		0.022
332	N20	H4 11-30	R19me2a	K20me1			0.031
333	N21	H4 11-30	R19me2a	K20me2			0.042
334	N22	H4 11-30	R19me2a	K20me3			0.037
335	N23	H4 11-30	R19me2a	K20ac			0.020
336	N24	H4 11-30	R19me2s	K20me1			0.025
337	O1	H4 11-30	R19me2s	K20me2			0.013
338	O2	H4 11-30	R19me2s	K20me3			0.010
339	O3	H4 11-30	R19me2s	K20ac			0.011
340	O4	H4 11-30	R24me2a	K20me1			0.040
341	O5	H4 11-30	R24me2a	K20me2			0.024
342	O6	H4 11-30	R24me2a	K20me3			0.029
343	O7	H4 11-30	R24me2a	K20ac			0.029
344	O8	H4 11-30	R24me2s	K20me1			0.071
345	O9	H4 11-30	R24me2s	K20me2			0.035
346	O10	H4 11-30	R24me2s	K20me3			0.025
347	O11	H4 11-30	R24me2s	K20ac			0.062
348	O12	H2a 1-19	unmod				1.000
349	O13	H2a 1-19	S1P				0.990
350	O14	H2a 1-19	K5ac				0.992
351	O15	H2a 1-19	K9ac				0.942
352	O16	H2a 1-19	K13ac				0.141

Appendix

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
353	O17	H2a 1-19	S1P	K5ac			0.999
354	O18	H2a 1-19	S1P	K9ac			1.000
355	O19	H2a 1-19	S1P	K13ac			0.995
356	O20	H2a 1-19	K5ac	K9ac			0.991
357	O21	H2a 1-19	K5ac	K13ac			0.999
358	O22	H2a 1-19	K9ac	K13ac			0.992
359	O23	H2a 1-19	S1P	K5ac	K9ac		0.975
360	O24	H2a 1-19	S1P	K5ac	K13ac		0.606
361	P1	H2a 1-19	S1P	K9ac	K13ac		0.634
362	P2	H2a 1-19	K5ac	K9ac	K13ac		0.059
363	P3	H2a 1-19	S1P	K5ac	K9ac	K13ac	0.894
364	P4	H2b 1-19	unmod				1.000
365	P5	H2b 1-19	K5ac				0.944
366	P6	H2b 1-19	K12ac				0.981
367	P7	H2b 1-19	S14P				0.994
368	P8	H2b 1-19	K15ac				0.939
369	P9	H2b 1-19	K5ac	K12ac			0.119
370	P10	H2b 1-19	K5ac	S14P			0.799
371	P11	H2b 1-19	K5ac	K15ac			0.899
372	P12	H2b 1-19	K12ac	S14P			0.248
373	P13	H2b 1-19	K12ac	K15ac			0.048
374	P14	H2b 1-19	S14P	K15ac			0.019
375	P15	H2b 1-19	K5ac	K12ac	S14P		0.925
376	P16	H2b 1-19	K5ac	K12ac	K15ac		0.044
377	P17	H2b 1-19	K5ac	S14P	K15ac		0.156
378	P18	H2b 1-19	K12ac	S14P	K15ac		0.040
379	P19	H2b 1-19	K5ac	K12ac	S14P	K15ac	0.017
380	P20		Biotin control				0.006
381	P21		c-Myc control				0.002
382	P22		Negative control				0.007
383	P23		Background 1				0.988
384	P24		Background 2				1.000

Pep. No.= Peptide number in one replica grid (left or right). Location= position of peptide on the grid. Mod= modified mark in the peptide. Norm. act.= normalized activity of spot intensity.



Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
1	A1	H3 1-19	unmod				0.896
2	A2	H3 1-19	R2me2s				0.744
3	A3	H3 1-19	R2me2a				0.755
4	A4	H3 1-19	R2Citr				0.618
5	A5	H3 1-19	T3P				0.499
6	A6	H3 1-19	K4me1				0.522
7	A7	H3 1-19	K4me2				0.229
8	A8	H3 1-19	K4me3				0.109
9	A9	H3 1-19	K4ac				0.299
10	A10	H3 1-19	R8me2s				0.192
11	A11	H3 1-19	R8me2a				0.451
12	A12	H3 1-19	R8Citr				0.368
13	A13	H3 1-19	K9me1				0.158
14	A14	H3 1-19	K9me2				0.107
15	A15	H3 1-19	K9me3				0.047
16	A16	H3 1-19	K9ac				0.018
17	A17	H3 1-19	S10P				0.011
18	A18	H3 1-19	T11P				0.011
19	A19	H3 1-19	K14ac				0.037
20	A20	H3 1-19	R2me2s	T3P			0.028
21	A21	H3 1-19	R2me2s	K4me1			0.012
22	A22	H3 1-19	R2me2s	K4me2			0.007
23	A23	H3 1-19	R2me2s	K4me3			0.012
24	A24	H3 1-19	R2me2s	K4ac			0.011
25	B1	H3 1-19	R2me2a	T3P			0.754
26	B2	H3 1-19	R2me2a	K4me1			0.541
27	B3	H3 1-19	R2me2a	K4me2			0.249
28	B4	H3 1-19	R2me2a	K4me3			0.220
29	B5	H3 1-19	R2me2a	K4ac			0.329
30	B6	H3 1-19	R2Citr	T3P			0.107
31	B7	H3 1-19	R2Citr	K4me1			0.053
32	B8	H3 1-19	R2Citr	K4me2			0.084
33	B9	H3 1-19	R2Citr	K4me3			0.079
34	B10	H3 1-19	R2Citr	K4ac			0.125
35	B11	H3 1-19	T3P	K4me1			0.670
36	B12	H3 1-19	T3P	K4me2			0.155
37	B13	H3 1-19	T3P	K4me3			0.128
38	B14	H3 1-19	T3P	K4ac			0.050
39	B15	H3 1-19	R2me2s	T3P	K4me1		0.015
40	B16	H3 1-19	R2me2s	T3P	K4me2		0.010
41	B17	H3 1-19	R2me2s	T3P	K4me3		0.031
42	B18	H3 1-19	R2me2s	T3P	K4ac		0.014
43	B19	H3 1-19	R2me2a	T3P	K4me1		0.031
44	B20	H3 1-19	R2me2a	T3P	K4me2		0.040
45	B21	H3 1-19	R2me2a	T3P	K4me3		0.027
46	B22	H3 1-19	R2me2a	T3P	K4ac		0.010
47	B23	H3 1-19	R8me2s	K9me1			0.001
48	B24	H3 1-19	R8me2s	K9me2			0.019
49	C1	H3 1-19	R8me2s	K9me3			0.287
50	C2	H3 1-19	R8me2s	K9ac			0.212
51	C3	H3 1-19	R8me2s	S10P			0.122
52	C4	H3 1-19	R8me2s	T11P			0.127
53	C5	H3 1-19	R8me2a	K9me1			0.104
54	C6	H3 1-19	R8me2a	K9me2			0.084
55	C7	H3 1-19	R8me2a	K9me3			0.027
56	C8	H3 1-19	R8me2a	K9ac			0.013
57	C9	H3 1-19	R8me2a	S10P			0.018
58	C10	H3 1-19	R8me2a	T11P			0.027
59	C11	H3 1-19	R8Citr	K9me1			0.085
60	C12	H3 1-19	R8Citr	K9me2			0.041
61	C13	H3 1-19	R8Citr	K9me3			0.043

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
62	C14	H3 1-19	R8Citr	K9ac			0.011
63	C15	H3 1-19	R8Citr	S10P			0.005
64	C16	H3 1-19	R8Citr	T11P			0.007
65	C17	H3 1-19	K9me1	S10P			0.015
66	C18	H3 1-19	K9me1	T11P			0.013
67	C19	H3 1-19	K9me1	K14ac			0.025
68	C20	H3 1-19	K9me2	S10P			0.020
69	C21	H3 1-19	K9me2	T11P			0.015
70	C22	H3 1-19	K9me2	K14ac			0.016
71	C23	H3 1-19	K9me3	S10P			0.004
72	C24	H3 1-19	K9me3	T11P			0.011
73	D1	H3 1-19	K9me3	K14ac			0.183
74	D2	H3 1-19	K9ac	S10P			0.069
75	D3	H3 1-19	K9ac	T11P			0.049
76	D4	H3 1-19	K9ac	K14ac			0.046
77	D5	H3 1-19	S10P	T11P			0.018
78	D6	H3 1-19	S10P	K14ac			0.019
79	D7	H3 1-19	T11P	K14ac			0.011
80	D8	H3 1-19	R8me2s	K9me1	S10P		0.012
81	D9	H3 1-19	R8me2s	K9me2	S10P		0.010
82	D10	H3 1-19	R8me2s	K9me3	S10P		0.011
83	D11	H3 1-19	R8me2s	K9ac	S10P		0.025
84	D12	H3 1-19	R8me2s	K9me1	T11P		0.027
85	D13	H3 1-19	R8me2s	K9me2	T11P		0.010
86	D14	H3 1-19	R8me2s	K9me3	T11P		0.007
87	D15	H3 1-19	R8me2s	K9ac	T11P		0.007
88	D16	H3 1-19	R8me2a	K9me1	S10P		0.007
89	D17	H3 1-19	R8me2a	K9me2	S10P		0.007
90	D18	H3 1-19	R8me2a	K9me3	S10P		0.017
91	D19	H3 1-19	R8me2a	K9ac	S10P		0.013
92	D20	H3 1-19	R8me2a	K9me1	T11P		0.024
93	D21	H3 1-19	R8me2a	K9me2	T11P		0.037
94	D22	H3 1-19	R8me2a	K9me3	T11P		0.022
95	D23	H3 1-19	R8me2a	K9ac	T11P		0.004
96	D24	H3 1-19	R8me2a	K9me1	S10P	T11P	0.073
97	E1	H3 1-19	R8me2a	K9me2	S10P	T11P	0.073
98	E2	H3 1-19	R8me2a	K9me3	S10P	T11P	0.068
99	E3	H3 1-19	R8me2a	K9ac	S10P	T11P	0.026
100	E4	H3 1-19	R2me2s	K4me1	R8me2s		0.012
101	E5	H3 1-19	R2me2s	K4me2	R8me2s		0.008
102	E6	H3 1-19	R2me2s	K4me3	R8me2s		0.007
103	E7	H3 1-19	R2me2s	K4ac	R8me2s		0.007
104	E8	H3 1-19	R2me2a	K4me1	R8me2a		0.005
105	E9	H3 1-19	R2me2a	K4me2	R8me2a		0.007
106	E10	H3 1-19	R2me2a	K4me3	R8me2a		0.011
107	E11	H3 1-19	R2me2a	K4ac	R8me2a		0.012
108	E12	H3 1-19	R2me2s	K4me1	K9me1		0.018
109	E13	H3 1-19	R2me2s	K4me2	K9me1		0.021
110	E14	H3 1-19	R2me2s	K4me3	K9me1		0.009
111	E15	H3 1-19	R2me2s	K4ac	K9me1		0.004
112	E16	H3 1-19	R2me2a	K4me1	K9me2		0.006
113	E17	H3 1-19	R2me2a	K4me2	K9me2		0.005
114	E18	H3 1-19	R2me2a	K4me3	K9me2		0.010
115	E19	H3 1-19	R2me2a	K4ac	K9me2		0.012
116	E20	H3 1-19	R2me2s	K4me1	K9me3		0.025
117	E21	H3 1-19	R2me2s	K4me2	K9me3		0.015
118	E22	H3 1-19	R2me2s	K4me3	K9me3		0.013
119	E23	H3 1-19	R2me2s	K4ac	K9me3		0.002
120	E24	H3 1-19	R2me2a	K4me1	K9ac		0.025
121	F1	H3 1-19	R2me2a	K4me2	K9ac		0.105
122	F2	H3 1-19	R2me2a	K4me3	K9ac		0.048
123	F3	H3 1-19	R2me2a	K4ac	K9ac		0.034
124	F4	H3 1-19	K4me1	R8me2s	K9me1		0.015
125	F5	H3 1-19	K4me2	R8me2s	K9me1		0.006
126	F6	H3 1-19	K4me3	R8me2s	K9me1		0.005
127	F7	H3 1-19	K4ac	R8me2s	K9me1		0.005
128	F8	H3 1-19	K4me1	R8me2a	K9me1		0.006
129	F9	H3 1-19	K4me2	R8me2a	K9me1		0.008
130	F10	H3 1-19	K4me3	R8me2a	K9me1		0.019
131	F11	H3 1-19	K4ac	R8me2a	K9me1		0.023
132	F12	H3 1-19	K4me1	R8me2s	K9me2		0.034
133	F13	H3 1-19	K4me2	R8me2s	K9me2		0.063
134	F14	H3 1-19	K4me3	R8me2s	K9me2		0.015
135	F15	H3 1-19	K4ac	R8me2s	K9me2		0.010
136	F16	H3 1-19	K4me1	R8me2a	K9me2		0.005
137	F17	H3 1-19	K4me2	R8me2a	K9me2		0.006

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
138	F18	H3 1-19	K4me3	R8me2a	K9me2		0.008
139	F19	H3 1-19	K4ac	R8me2a	K9me2		0.010
140	F20	H3 1-19	K4me1	R8me2s	K9me3		0.013
141	F21	H3 1-19	K4me2	R8me2s	K9me3		0.010
142	F22	H3 1-19	K4me3	R8me2s	K9me3		0.009
143	F23	H3 1-19	K4ac	R8me2s	K9me3		0.013
144	F24	H3 1-19	K4me1	R8me2a	K9me3		0.022
145	G1	H3 1-19	K4me2	R8me2a	K9me3		0.211
146	G2	H3 1-19	K4me3	R8me2a	K9me3		0.100
147	G3	H3 1-19	K4ac	R8me2a	K9me3		0.054
148	G4	H3 1-19	K4me1	R8me2s	K9ac		0.030
149	G5	H3 1-19	K4me2	R8me2s	K9ac		0.014
150	G6	H3 1-19	K4me3	R8me2s	K9ac		0.008
151	G7	H3 1-19	K4ac	R8me2s	K9ac		0.004
152	G8	H3 1-19	K4me1	R8me2a	K9ac		0.004
153	G9	H3 1-19	K4me2	R8me2a	K9ac		0.009
154	G10	H3 1-19	K4me3	R8me2a	K9ac		0.059
155	G11	H3 1-19	K4ac	R8me2a	K9ac		0.051
156	G12	H3 1-19	R2me2s	K4me1	R8me2s	K9me1	0.050
157	G13	H3 1-19	R2me2s	K4me2	R8me2s	K9me1	0.067
158	G14	H3 1-19	R2me2s	K4me3	R8me2s	K9me1	0.021
159	G15	H3 1-19	R2me2s	K4ac	R8me2s	K9me1	0.011
160	G16	H3 1-19	R2me2a	K4me1	R8me2s	K9me1	0.009
161	G17	H3 1-19	R2me2a	K4me2	R8me2s	K9me1	0.010
162	G18	H3 1-19	R2me2a	K4me3	R8me2s	K9me1	0.010
163	G19	H3 1-19	R2me2a	K4ac	R8me2s	K9me1	0.007
164	G20	H3 1-19	R2me2s	K4me1	R8me2s	K9me2	0.004
165	G21	H3 1-19	R2me2s	K4me2	R8me2s	K9me2	0.005
166	G22	H3 1-19	R2me2s	K4me3	R8me2s	K9me2	0.008
167	G23	H3 1-19	R2me2s	K4ac	R8me2s	K9me2	0.017
168	G24	H3 1-19	R2me2a	K4me1	R8me2s	K9me2	0.018
169	H1	H3 1-19	R2me2a	K4me2	R8me2s	K9me2	0.402
170	H2	H3 1-19	R2me2a	K4me3	R8me2s	K9me2	0.257
171	H3	H3 1-19	R2me2a	K4ac	R8me2s	K9me2	0.092
172	H4	H3 1-19	R2me2s	K4me1	R8me2s	K9me3	0.103
173	H5	H3 1-19	R2me2s	K4me2	R8me2s	K9me3	0.025
174	H6	H3 1-19	R2me2s	K4me3	R8me2s	K9me3	0.010
175	H7	H3 1-19	R2me2s	K4ac	R8me2s	K9me3	0.004
176	H8	H3 1-19	R2me2a	K4me1	R8me2s	K9me3	0.005
177	H9	H3 1-19	R2me2a	K4me2	R8me2s	K9me3	0.011
178	H10	H3 1-19	R2me2a	K4me3	R8me2s	K9me3	0.026
179	H11	H3 1-19	R2me2a	K4ac	R8me2s	K9me3	0.046
180	H12	H3 1-19	R2me2s	K4me1	R8me2s	K9ac	0.105
181	H13	H3 1-19	R2me2s	K4me2	R8me2s	K9ac	0.052
182	H14	H3 1-19	R2me2s	K4me3	R8me2s	K9ac	0.019
183	H15	H3 1-19	R2me2s	K4ac	R8me2s	K9ac	0.010
184	H16	H3 1-19	R2me2a	K4me1	R8me2s	K9ac	0.013
185	H17	H3 1-19	R2me2a	K4me2	R8me2s	K9ac	0.027
186	H18	H3 1-19	R2me2a	K4me3	R8me2s	K9ac	0.012
187	H19	H3 1-19	R2me2a	K4ac	R8me2s	K9ac	0.012
188	H20	H3 1-19	R2me2s	K4me1	R8me2a	K9me1	0.005
189	H21	H3 1-19	R2me2s	K4me2	R8me2a	K9me1	0.006
190	H22	H3 1-19	R2me2s	K4me3	R8me2a	K9me1	0.017
191	H23	H3 1-19	R2me2s	K4ac	R8me2a	K9me1	0.031
192	H24	H3 1-19	R2me2a	K4me1	R8me2a	K9me1	0.004
193	I1	H3 1-19	R2me2a	K4me2	R8me2a	K9me1	0.677
194	I2	H3 1-19	R2me2a	K4me3	R8me2a	K9me1	0.517
195	I3	H3 1-19	R2me2a	K4ac	R8me2a	K9me1	0.297
196	I4	H3 1-19	R2me2s	K4me1	R8me2a	K9me2	0.059
197	I5	H3 1-19	R2me2s	K4me2	R8me2a	K9me2	0.029
198	I6	H3 1-19	R2me2s	K4me3	R8me2a	K9me2	0.008
199	I7	H3 1-19	R2me2s	K4ac	R8me2a	K9me2	0.005
200	I8	H3 1-19	R2me2a	K4me1	R8me2a	K9me2	0.006
201	I9	H3 1-19	R2me2a	K4me2	R8me2a	K9me2	0.014
202	I10	H3 1-19	R2me2a	K4me3	R8me2a	K9me2	0.023
203	I11	H3 1-19	R2me2a	K4ac	R8me2a	K9me2	0.021
204	I12	H3 1-19	R2me2s	K4me1	R8me2a	K9me3	0.020
205	I13	H3 1-19	R2me2s	K4me2	R8me2a	K9me3	0.016
206	I14	H3 1-19	R2me2s	K4me3	R8me2a	K9me3	0.011
207	I15	H3 1-19	R2me2s	K4ac	R8me2a	K9me3	0.008
208	I16	H3 1-19	R2me2a	K4me1	R8me2a	K9me3	0.018
209	I17	H3 1-19	R2me2a	K4me2	R8me2a	K9me3	0.054
210	I18	H3 1-19	R2me2a	K4me3	R8me2a	K9me3	0.060
211	I19	H3 1-19	R2me2a	K4ac	R8me2a	K9me3	0.025
212	I20	H3 1-19	R2me2s	K4me1	R8me2a	K9ac	0.014
213	I21	H3 1-19	R2me2s	K4me2	R8me2a	K9ac	0.010

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
214	I22	H3 1-19	R2me2s	K4me3	R8me2a	K9ac	0.020
215	I23	H3 1-19	R2me2s	K4ac	R8me2a	K9ac	0.111
216	I24	H3 1-19	R2me2a	K4me1	R8me2a	K9ac	0.192
217	J1	H3 1-19	R2me2a	K4me2	R8me2a	K9ac	0.822
218	J2	H3 1-19	R2me2a	K4me3	R8me2a	K9ac	0.622
219	J3	H3 1-19	R2me2a	K4ac	R8me2a	K9ac	0.201
220	J4	H3 7-26	unmod				0.168
221	J5	H3 7-26	K14ac				0.039
222	J6	H3 7-26	K14ac	S10P			0.014
223	J7	H3 7-26	K14ac	T11P			0.008
224	J8	H3 7-26	R17me2s				0.009
225	J9	H3 7-26	R17me2a				0.009
226	J10	H3 7-26	R17Citr				0.013
227	J11	H3 7-26	K18ac				0.024
228	J12	H3 7-26	K14ac	R17me2s			0.015
229	J13	H3 7-26	K14ac	R17me2a			0.023
230	J14	H3 7-26	K14ac	K18ac			0.017
231	J15	H3 7-26	R17me2s	K18ac			0.012
232	J16	H3 7-26	R17me2a	K18ac			0.039
233	J17	H3 7-26	R17Citr	K18ac			0.098
234	J18	H3 7-26	K14ac	R17me2s	K18ac		0.057
235	J19	H3 7-26	K14ac	R17me2a	K18ac		0.263
236	J20	H3 16-35	unmod				0.049
237	J21	H3 16-35	R26me2s				0.037
238	J22	H3 16-35	R26me2a				0.052
239	J23	H3 16-35	R26Citr				0.051
240	J24	H3 16-35	K27me1				0.342
241	K1	H3 16-35	K27me2				0.862
242	K2	H3 16-35	K27me3				0.573
243	K3	H3 16-35	K27ac				0.248
244	K4	H3 16-35	S28P				0.019
245	K5	H3 16-35	R26me2s	K27me1			0.132
246	K6	H3 16-35	R26me2s	K27me2			0.017
247	K7	H3 16-35	R26me2s	K27me3			0.011
248	K8	H3 16-35	R26me2s	K27ac			0.004
249	K9	H3 16-35	R26me2s	S28P			0.004
250	K10	H3 16-35	R26me2a	K27me1			0.007
251	K11	H3 16-35	R26me2a	K27me2			0.010
252	K12	H3 16-35	R26me2a	K27me3			0.004
253	K13	H3 16-35	R26me2a	K27ac			0.005
254	K14	H3 16-35	R26me2a	S28P			0.000
255	K15	H3 16-35	R26Citr	K27me1			0.011
256	K16	H3 16-35	R26Citr	K27me2			0.042
257	K17	H3 16-35	R26Citr	K27me3			0.036
258	K18	H3 16-35	R26Citr	S28P			0.004
259	K19	H3 16-35	K27me1	S28P			0.004
260	K20	H3 16-35	K27me2	S28P			0.004
261	K21	H3 16-35	K27me3	S28P			0.004
262	K22	H3 16-35	K27ac	S28P			0.004
263	K23	H3 16-35	R26me2s	K27me1	S28P		0.015
264	K24	H3 16-35	R26me2s	K27me2	S28P		0.022
265	L1	H3 16-35	R26me2s	K27me3	S28P		0.166
266	L2	H3 16-35	R26me2s	K27ac	S28P		0.091
267	L3	H3 16-35	R26me2a	K27me1	S28P		0.044
268	L4	H3 16-35	R26me2a	K27me2	S28P		0.024
269	L5	H3 16-35	R26me2a	K27me3	S28P		0.012
270	L6	H3 16-35	R26me2a	K27ac	S28P		0.008
271	L7	H3 26-45	unmod				0.038
272	L8	H3 26-45	K36me1				0.011
273	L9	H3 26-45	K36me2				0.020
274	L10	H3 26-45	K36me3				0.006
275	L11	H3 26-45	K36ac				0.011
276	L12	H4 1-19	unmod				0.004
277	L13	H4 1-19	S1P				0.007
278	L14	H4 1-19	R3me2s				0.011
279	L15	H4 1-19	R3me2a				0.030
280	L16	H4 1-19	K5ac				0.107
281	L17	H4 1-19	K8ac				0.444
282	L18	H4 1-19	K12ac				0.187
283	L19	H4 1-19	K16ac				0.050
284	L20	H4 1-19	S1P	R3me2s			0.015
285	L21	H4 1-19	S1P	R3me2a			0.010
286	L22	H4 1-19	S1P	K5ac			0.032
287	L23	H4 1-19	R3me2s	K5ac			0.172
288	L24	H4 1-19	R3me2s	K8ac			0.222
289	M1	H4 1-19	R3me2a	K5ac			0.750

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
290	M2	H4 1-19	R3me2a	K8ac			0.594
291	M3	H4 1-19	K5ac	K8ac			0.300
292	M4	H4 1-19	K8ac	K12ac			0.111
293	M5	H4 1-19	K8ac	K16ac			0.194
294	M6	H4 1-19	K12ac	K16ac			0.048
295	M7	H4 1-19	S1P	R3me2s	K5ac		0.019
296	M8	H4 1-19	S1P	R3me2a	K5ac		0.012
297	M9	H4 1-19	R3me2s	K5ac	K8ac		0.015
298	M10	H4 1-19	R3me2a	K5ac	K8ac		0.013
299	M11	H4 1-19	K5ac	K8ac	K12ac		0.014
300	M12	H4 1-19	K8ac	K12ac	K16ac		0.014
301	M13	H4 1-19	S1P	R3me2s	K5ac	K8ac	0.017
302	M14	H4 1-19	S1P	R3me2a	K5ac	K8ac	0.018
303	M15	H4 1-19	R3me2s	K5ac	K8ac	K12ac	0.329
304	M16	H4 1-19	R3me2a	K5ac	K8ac	K12ac	0.346
305	M17	H4 1-19	K5ac	K8ac	K12ac	K16ac	0.521
306	M18	H4 11-30	unmod				0.435
307	M19	H4 11-30	K12ac				0.238
308	M20	H4 11-30	K16ac				0.233
309	M21	H4 11-30	R17me2s				0.027
310	M22	H4 11-30	R17me2a				0.034
311	M23	H4 11-30	R19me2s				0.157
312	M24	H4 11-30	R19me2a				0.186
313	N1	H4 11-30	K20me1				0.968
314	N2	H4 11-30	K20me2				0.711
315	N3	H4 11-30	K20me3				0.465
316	N4	H4 11-30	K20ac				0.271
317	N5	H4 11-30	R24me2a				0.264
318	N6	H4 11-30	R24me2s				0.071
319	N7	H4 11-30	K12ac	K16ac			0.052
320	N8	H4 11-30	K16ac	R17me2s			0.057
321	N9	H4 11-30	K16ac	R17me2a			0.128
322	N10	H4 11-30	K16ac	R19me2s			0.090
323	N11	H4 11-30	K16ac	R19me2a			0.019
324	N12	H4 11-30	K16ac	K20me1			0.040
325	N13	H4 11-30	K16ac	K20me2			0.203
326	N14	H4 11-30	K16ac	K20me3			0.373
327	N15	H4 11-30	K16ac	K20ac			0.360
328	N16	H4 11-30	K12ac	K16ac	K20me1		0.625
329	N17	H4 11-30	K12ac	K16ac	K20me2		0.480
330	N18	H4 11-30	K12ac	K16ac	K20me3		0.079
331	N19	H4 11-30	K12ac	K16ac	K20ac		0.010
332	N20	H4 11-30	R19me2a	K20me1			0.063
333	N21	H4 11-30	R19me2a	K20me2			0.037
334	N22	H4 11-30	R19me2a	K20me3			0.039
335	N23	H4 11-30	R19me2a	K20ac			0.081
336	N24	H4 11-30	R19me2s	K20me1			0.092
337	O1	H4 11-30	R19me2s	K20me2			0.951
338	O2	H4 11-30	R19me2s	K20me3			0.839
339	O3	H4 11-30	R19me2s	K20ac			0.608
340	O4	H4 11-30	R24me2a	K20me1			0.736
341	O5	H4 11-30	R24me2a	K20me2			0.449
342	O6	H4 11-30	R24me2a	K20me3			0.168
343	O7	H4 11-30	R24me2a	K20ac			0.100
344	O8	H4 11-30	R24me2s	K20me1			0.149
345	O9	H4 11-30	R24me2s	K20me2			0.096
346	O10	H4 11-30	R24me2s	K20me3			0.143
347	O11	H4 11-30	R24me2s	K20ac			0.082
348	O12	H2a 1-19	unmod				0.051
349	O13	H2a 1-19	S1P				0.240
350	O14	H2a 1-19	K5ac				0.440
351	O15	H2a 1-19	K9ac				0.522
352	O16	H2a 1-19	K13ac				0.808
353	O17	H2a 1-19	S1P	K5ac			0.711
354	O18	H2a 1-19	S1P	K9ac			0.606
355	O19	H2a 1-19	S1P	K13ac			0.098
356	O20	H2a 1-19	K5ac	K9ac			0.044
357	O21	H2a 1-19	K5ac	K13ac			0.014
358	O22	H2a 1-19	K9ac	K13ac			0.190
359	O23	H2a 1-19	S1P	K5ac	K9ac		0.050
360	O24	H2a 1-19	S1P	K5ac	K13ac		0.094
361	P1	H2a 1-19	S1P	K9ac	K13ac		1.000
362	P2	H2a 1-19	K5ac	K9ac	K13ac		0.972
363	P3	H2a 1-19	S1P	K5ac	K9ac	K13ac	0.868
364	P4	H2b 1-19	unmod				0.799
365	P5	H2b 1-19	K5ac				0.733

Appendix

Peptide No.	Location	Name	Mod 1	Mod 2	Mod 3	Mod 4	Norm. act.
366	P6	H2b 1-19	K12ac				0.585
367	P7	H2b 1-19	S14P				0.591
368	P8	H2b 1-19	K15ac				0.387
369	P9	H2b 1-19	K5ac	K12ac			0.153
370	P10	H2b 1-19	K5ac	S14P			0.127
371	P11	H2b 1-19	K5ac	K15ac			0.025
372	P12	H2b 1-19	K12ac	S14P			0.026
373	P13	H2b 1-19	K12ac	K15ac			0.022
374	P14	H2b 1-19	S14P	K15ac			0.023
375	P15	H2b 1-19	K5ac	K12ac	S14P		0.817
376	P16	H2b 1-19	K5ac	K12ac	K15ac		0.812
377	P17	H2b 1-19	K5ac	S14P	K15ac		0.112
378	P18	H2b 1-19	K12ac	S14P	K15ac		0.026
379	P19	H2b 1-19	K5ac	K12ac	S14P	K15ac	0.017
380	P20		Biotin control				0.004
381	P21		c-Myc control				0.004
382	P22		Negative control				0.004
383	P23		Background 1				0.073
384	P24		Background 2				0.038

Pep. No.= Peptide number in one replica grid (left or right). Location= position of peptide on the grid. Mod= modified mark in the peptide. Norm. act.= normalized activity of spot intensity.